

Who should call the shots?

Britain's environment research council has an impossible burden and an inadequate budget. It should return responsibility where it belongs, with politicians.

THE management of a growing scientific enterprise is child's play compared with the management of a shrinking enterprise. That, charity demands, is the spirit in which even sceptics should read the prospectus published this week by Britain's newest and smallest research council, the Natural Environment Research Council (see pp. 516 and 523). The kind thing to say about the council's present difficulty is that it never had a fair chance. The intention, twenty years ago, was that the council should be a comprehensive umbrella for the conduct of research in all aspects of the natural environment, but this was frustrated by (among other things) the political decision that the Meteorological Office should not be counted in. Since then, the council has been financially and intellectually damaged by the unwillingness of central government departments to commission the volume of research foreseen at the time of the Rothschild reorganization of British civil science just over a decade ago. And there have been self-inflicted wounds, those caused by poor organization for example.

The new plan is an attempt at repair and is, in many ways, courageous. Importing the doctrines of management responsibility into the affairs of a research organization is bound to cause trouble, notably among those who will no longer be flattered by being asked to give advice but not to face the consequences thereof. Institute directors will also make trouble, sometimes on the good grounds that they know better than the committees what work should be done, sometimes no doubt out of mere pique. There will also be resentment, much of it justifiable, among those researchers who will be forced (or tempted) away from their jobs for the sake of releasing funds for the support of academic research. The council deserves high marks for bravery in having grasped these nettles. But is the task on which it has embarked either necessary or prudent?

Reorganization

The chief and laudable justification of the reorganization now proposed is to shift the balance of the council's research support from its own institutes to the universities, but the sums of money involved are only small. On the council's own figures, spending on research grants and training awards will increase from £11.6 million in the present year to £16.9 million five years from now. With British academic research in the doldrums, the extra £5 million would be eagerly welcomed if it were available now; what joy it will bring five years from now is anybody's guess. But the council's spending on this valuable cause is hardly ever likely to exceed ten per cent of what is spent in the universities by the research councils for science (with engineering) and medicine. When the scale of the environment council's operations is, by comparison, so much smaller, the advantages of pluralism are likely to be more than offset by the danger that the council's support of academic research will become the prisoner of sectional interests. It would make more sense that the council's academic support programme should be made part of the larger pot operated by the Science and Engineering Research Council.

The second part of the council's prospectus entails further beastliness towards its own research institutes, most of which were inherited in 1967 much as they still are. The institutes are a mixed bag. Most are important sources of research in their fields, but some serve also as important means for supporting academic

research. Others, such as the Institute of Terrestrial Ecology, have both these functions and are also the natural source of intellectual support for bodies such as the Nature Conservancy Council (which is desperately in need of a research organization to call its own). Still others have important executive functions; the British Antarctic Survey has become almost an instrument of British foreign policy since the Falklands War (spending nearly 10 per cent of the council's budget), but the British Geological Survey, which earns two-thirds of the council's income from research commissions but whose net cost is much the same, could be held to satisfy a national economic need. The obvious danger in the research council's proposals is that the effectiveness of some of these institutes could be further damaged by the pressure on their resources without there ever having been an objective appraisal of their present value to the intellectual community or to the British national interest. It will be especially cruel (and wasteful) if this should happen simply because there has been a sequence of arbitrary budget cuts now further to be intensified by the environment research council's laudable ambition to do more directly to support academic research. Why, at least, not preface such a decision by a protest to the paymasters (the Advisory Board for the Research Councils and, ultimately, the British government) that the casualties may be vital research institutions?

Management

What the new prospectus says about the management of its operations is bound to cause trouble, but could be worth a try. The essence of the scheme is a greater degree of central authority for the shaping of research programmes, which is for obvious reasons an unfamiliar notion. Those who emphasize the anarchic character of research will be up in arms, but the more immediate difficulty stems from the incompleteness and diversity of the research council's operations, which are as lopsided now as when it began. The absence of the Meteorological Office means that the council's research in the atmospheric sciences is hardly as important as that carried out by other organizations, the Atomic Energy Authority for example. The lack of a formal connection with the Nature Conservancy Council harms both partners in this unrequited relationship. The British Antarctic Survey is beyond appraisal (which is no reflection on the quality of its work). The research council has ambitions in remote sensing from beyond the atmosphere, but space research is managed by others. Taking a central view in such circumstances will be difficult, to say the least. The council's brave administrative experiment had best been attempted in some other field.

These are some of the reasons why *Nature* argued, last October (311, 498; 1984), that the best solution of the council's problem would be its dismemberment. Mr Hugh Fish, the council's chairman, demurred, saying that the option had been considered but rejected and that many of the council's establishments do excellent work, which is true (see *Nature* 311, 700; 1984). But if and when the council is reorganized along the lines now proposed, it will still be a lopsided organization and, given the financial forecasts, still less able than at present to perform the functions expected of it. No doubt a part of the council's interest that there should be a close examination of what its founding charter means is that it would be a great convenience if some of



its functions could be dropped altogether, but that is an invidious position in which to be discovered. Even the environment council's detractors will agree that it is grossly unfair that a public corporation managed by people who are not even civil servants should be saddled with impossible responsibilities and an inadequate budget and invited to make decisions which are, in essence, political. Whether and on what scale there should be a continuing geological survey in Britain is a decision that constitutionally should be made by ministers, not by research council employees or, still less, by volunteer members of advisory committees such as the Advisory Board for the Research Councils. For the time being, the research council's duty is to tell them so, in terms that cannot be shirked. Doing so would necessarily invite attention to the research council's own plight, as it would become if the British Geological Survey were to join the British Antarctic Survey as a protected pensioner of the "protected" British science budget. Ministerial inquirers might well conclude that for the research council, dismemberment would certainly then be the best policy. People are naturally shy of opening that can of worms, but, in propriety, there is no choice. □

Star wars bargain

The US president's science adviser should learn to temper his loyalty with proper embarrassment.

EVEN under the best of circumstances, one of the essential requirements for the job of science adviser to the US president is a split personality. The scientific community sees the science adviser as a superlobbyist for science — but he is a lobbyist who is also part of the political apparatus, which makes for strain when the apparatus takes decisions not necessarily in the best interests of science. A tight budget is the acid test. Dr George Keyworth's predecessor was able to retain sufficient aloofness from the land of make-believe of Washington politics to appear embarrassed, at least, when forced during the 1980 election to make a series of budgetary about-turns by President Carter seem to be a sudden large new investment in research and development. Keyworth, however, is another matter. He has never made any bones about his conviction that his primary job is to support the president. Keyworth has shown no trace of embarrassment about star wars in the past two years and, last week, in announcing how the 1986 budget will affect science, he showed no embarrassment in playing a little fast and loose with the figures to make it look better than it was. Among his more incorrect statements was that "funds obligated for non-defense research and development will remain essentially constant from 1985 to 1986". The budget documents disagree: they show a 4 per cent cut in non-defense research and development and that on top of a 4 per cent or so inflation rate. It is hard to see how an 8 per cent real cut in spending is "essentially" no change.

Keyworth played a more standard Washington game, though one that is no more defensible, in claiming that funds for alternatives to the UNESCO science programmes (see p.518) were to be found in the budgets for various US science agencies. In fact, there is no new money at all except for \$2.7 million allowed by the State Department (a far cry from the \$50 million that the United States had been contributing to UNESCO and even from the \$14 million that the National Academy of Sciences recommended be provided for science activities to make up for the US withdrawal). And while it is true that the National Science Foundation supports a relatively large amount of international cooperative research, this has been going on for years and has nothing to do with UNESCO or the decision to withdraw therefrom. And in any case, international programmes have been cut in the foundation's budget as now proposed. The same game was played in claiming the existence of \$42 million in the Department of Agriculture's Cooperative State Research Service (CSRS) for biotechnology. This is not new money; the CSRS budget has been frozen.

In testimony before the House of Representatives' Science and Technology Committee last week, Keyworth also glibly defended

the administration decision to cut federal assistance to students, noting that parents are demonstrating a new sense of responsibility to educate their children. He has defended the budget's indifference to US-Soviet scientific exchanges, saying, in response to a congressman's admonition to keep an open mind on the subject, that "I'm not sure that I would want to have a terribly open mind in dealing with a country that has an inherently adversarial form of government". He defended the administration request of \$3,700 million for star wars: "The Soviets have more than one reason to worry about SDI [the Strategic Defense Initiative] in addition to the prospect that strategic defense will make worthless the bulk of their immense offensive nuclear force. Our effort to produce a defensive system will almost certainly stimulate a new thrust in technology and leave the stagnating Soviet economy further behind."

Keyworth summed up by saying, "I'm proud of the budget". It would have been reassuring to the scientific community if he had at least looked a bit embarrassed when he said it. □

Loyal Keyworth

The United States cannot expect next month's Geneva talks to succeed without a bargain.

MR Caspar Weinberger's swing through Western Europe earlier this week is the most serious attempt so far by the US administration to win support for star wars, otherwise the Strategic Defense Initiative (SDI). Chancellor Helmut Kohl in West Germany is Mr Weinberger's most conspicuous convert so far, although the visit of Mrs Margaret Thatcher, the British prime minister, to Washington at the end of last year leaves her on record as having won from President Reagan an assurance that SDI is at this stage strictly a research programme, one that will not be turned into a working military system without consultations on the subject with the Soviet Union. Mr Weinberger might tactfully have made more of this important concession in his declarations at the weekend, especially because much of Western Europe is fearful that the talks arranged for Geneva on 12 March between the Soviet Union and the United States will collapse in the warm glow of the belief engendered by President Reagan's inaugural speech that all is well in the West. Neither Mr Weinberger nor the US administration seems to appreciate that the pursuit of SDI requires some kind of tacit licence from the Soviet Union, at least if the arms control negotiations are to survive.

If SDI is indeed a research programme, the following considerations apply. First, it is by no means certain that the outcome will be a workable defence against strategic missiles. Indeed, there are many who insist that the opposite must be the case, and that the only conceivable outcome of the efforts now being expended will be a kind of improved version of the anti-ballistic missile systems canvassed in the early 1970s and rejected then on the grounds that they would be ruinously expensive and strategically destabilizing. So neither the United States nor, for that matter, West Germany, should count on success, which means that the need for an agreement (if attainable) at Geneva is as urgent as it was before SDI became fashionable. Second, then, it follows that there should be some demonstration at Geneva that SDI is indeed pure exploration, for which purpose the most convenient instrument would be a moratorium (say, for the next four years of the present administration) on the testing of anti-satellite weapons. The case, repeated last week by Mr Weinberger, that the Soviet Union has already deployed such a weapons system is unconvincing, for it is too primitive a system to be taken seriously at this stage. Third, there is the problem of the anti-ballistic missile treaty, which formally prohibits the development and testing of components of such a system beyond the atmosphere. No doubt there are lawyers who will argue that the bits and pieces that will be developed as part of SDI are not components of an anti-ballistic missile system, for such a device has not yet been designed. But there are lawyers, many of them from the Soviet Union, who will argue the other way. The only way of finessing that difficulty is to have it openly discussed at Geneva. □

US shuttle

NASA takes gloomy view of prospects

Washington

COMPETITION between the United States and Europe to launch commercial payloads into Earth orbit is apparently bruising the National Aeronautics and Space Administration (NASA).

Administrator James Beggs of NASA told Congress last week that market prospects for the "Space Transportation System" (NASAspeak for the shuttle) has been "very severely eroded" in the past year, so much so that NASA is having difficulty justifying a fifth orbiter in the foreseeable future.

Competition from Europe is not the only threat to the shuttle's commercial prospects but it is one of the most worrying to NASA. Beggs told the House of Representatives Science and Technology Committee that Europe was marketing the Ariane launch vehicle "very aggressively" and said he was "pessimistic about our ability to hold the lion's share of the market". In the past 14 months, Beggs said, "they've won five and we've won five and nobody else has won any". He also noted the recent decision of European science ministers to commit "around 1 billion" dollars to developing the Ariane 4 and Ariane 5 launchers.

But there are other problems for NASA's shuttle commercialization plans that stem from nearer home. Last year the Department of Defense decided to go ahead with its own expendable launch vehicle for military payloads, partly because it is disillusioned with the shuttle's reliability (numerous launches have had to be postponed) but also because of publicity surrounding shuttle launches. Press coverage of a military shuttle mission last month caused a political furor because Secretary of Defense Caspar Weinberger had said that news media must not so much as speculate about the objectives of the mission.

NASA had been expecting to launch at least two military payloads a year, roughly 10 per cent of planned missions for the next few years, and the loss of that source of income will hurt. Moreover, the National Oceanic and Atmospheric Administration (NOAA) has recently informed NASA that three meteorological satellites scheduled for launch in the shuttle will now be sent up on reconditioned Titan II rockets which have for years been sitting unused in silos as Minuteman missile launchers. Beggs noted dryly that the reconditioned rockets "are being offered around town" at cut-price rates but declined to comment on the financial wisdom of NOAA's decision to abandon the shuttle and to use its own launchers.

Beggs said that by the early 1990s the country would probably be ready to start developing a second generation shuttle. Enough has already been learned for clear technological improvements to be incorporated, such as an improved thermal insulation system. Problems with the shuttle's heat-resistant tiles have caused extensive delays to launch schedules. A second generation shuttle might be ready for operation early in the next century, Beggs said. NASA's proposed budget for 1986 (\$7,900 million) will not mean the cancellation of any programmes, but development of the space station in 1986 will have to proceed at a "slightly slower pace" than had been anticipated, and no new projects will be started.

The Ariane 4 rocket now under development by the European Space Agency will be able to take a payload of around 4 metric tonnes into geosynchronous orbit; Ariane 5, which is still on the drawing board, would double that performance. In comparison, a shuttle payload can be up

Ariane up again

THE twelfth Ariane launch last Friday, using an Ariane-3 rocket, successfully placed two communications satellites into geostationary transfer orbit: Arabsat F1 (1,195 kg launch mass) and Brasilsat 1 (1,140 kg). Arabsat will serve North Africa and the Middle East, using ground control stations at Dhihrab (near Riyadh) and near Tunis. A second Arabsat is due to be launched later this year by the space shuttle. Brasilsat, naturally enough, will serve as a national Brazilian communications satellite, and will be controlled by the Tanguá ground station near Rio de Janeiro. The satellites are firsts both for the Arab states and for Brazil. Robert Walgate

to 30 tonnes, although extra motors are needed to boost a satellite from the shuttle's low orbit into geosynchronous orbit.

At present, NASA is flying experimental packages in the shuttle for potential commercial collaborators at bargain basement prices (some are free), but NASA has been directed to move towards a realistic pricing policy for commercial shuttle customers. As it does so, Ariane can be expected to enjoy an increasing cost advantage over the shuttle. At least, that is what NASA people fear and what Europeans hope.

Tim Beardsley

Academics' doubts on research plan

BRITISH academics familiar with the Natural Environment Research Council (NERC) say that the corporate plan contains good news (more emphasis on university support) and bad news (the continuing problem of declining funds). Two issues emerge as particularly contentious, the increased role of central administration and, among the Earth scientists, the apparent lack of commitment to the British Geological Survey.

Under the new plan, due to be published today (14 February) the role of central administrative staff would be enhanced in three ways. The number of such staff serving on the scientific committees would be increased, three scientific directors (possibly permanent) would be appointed to headquarters and the position of the NERC secretary as the conduit between the chairman of NERC and those below him would be confirmed. All three developments ring alarm bells. "I would be horrified to think that the scientific directors might be permanent members of staff — I cannot think of anyone who could do the job", said one academic. Another feared that the chairman of the research council would be insulated from the directors of science. And there are complaints of "occasions when the lack of administrative staff on the scientific committees has produced problems of coordination, but to give them voting rights certainly enhances the degree to which control is centralized". One com-

mon view is that the scientific directors could be an improvement, but that "anyone doing the job for more than four or five years could become a danger to the community".

The recent problems (see *Nature* 311, 499; 1984) of the systematic geological surveying by the British Geological Survey (BGS) have stimulated much public criticism of NERC by British academics. This concern is recognized but left virtually unanswered in the corporate plan. "Certainly the structure of BGS needs changing and contraction", said one senior geologist, "but what is desperately needed is some commitment to a systematic survey of the United Kingdom, both at the surface and at depth". The corporate plan promises "alternative approaches, with a concomitant loss of precision, if additional resources cannot be found". The document says elsewhere that "the council is reviewing its role of support of routine survey and data banking exercises: the latter will be pursued only on commission and/or where they underpin major programmes of research". That members of BGS are critical of their parent body, NERC, has become well known (see *Nature* 31 January, p.340). But there is now an influential body of academic opinion that believes that the strategic geological survey of the United Kingdom should be removed from NERC's control, perhaps to a government department.

Philip Campbell

UNESCO

US goes back on promise

Washington

IN spite of a public commitment to continue supporting the science programmes of UNESCO (the United Nations Educational, Scientific and Cultural Organization), or their equivalent through other channels, after the United States' resignation from the international organization with effect from 1 January this year, the Reagan administration last week requested only a token \$2.7 million in the budget it sent to Congress. The US contribution to UNESCO science programmes had been \$14 million per year, according to an analysis prepared last year by the National Academy of Sciences.

The commitment had been offered most recently by Assistant Secretary of State Gregory Newell in testimony before the

House of Representatives Foreign Affairs Committee last autumn. It is thought that the State Department's original budget request included the money for the science programme, but was shot down by the Office of Management and Budget.

State Department officials say they do not yet know exactly how the \$2.7 million will be allocated. But some of the money is earmarked for non-science activities, such as continued US participation in the international copyright convention. The \$2 million or so that will be left for science is expected to provide for minimal continued US participation in five UNESCO programmes: the Intergovernmental Oceanographic Commission (IOC), the International Geological Correlation Program (IGCP), the International Hydrological Program (IHP), the Man and Biosphere (MAB) programme and the natural hazards programme. A State Department official said that the aim will be to "plug the holes where US participation is essential to US interests".

But the State Department has doubts about how far \$2 million will go to meet that aim. Although there is no dispute that the United States can retain its seat on IOC, and possibly on the governing board of IGCP, it may lose influence if it does not contribute significantly to overhead costs and to new projects. In the other programmes, the United States is sure to lose its official role in management and planning. And even the \$2 million, designed to support directly the cost of participation by US scientists plus a share of overhead, falls short of the bare minimum figure of \$3 million recommended by the National Academy of Sciences report for direct contribution to these five UNESCO programmes.

The academy, moreover, recommended substantial additional support through alternative channels to make up for the US withdrawal. No such new funds are provided in the proposed 1986 budget; in fact, the National Science Foundation's budget for international cooperative research projects is to be cut from \$13.6 to \$12.1 million under the Reagan proposal, reducing participation by US scientists to 1,300 from the present 1,400.

The academy report warned that "in order to maintain confidence in US participation in international science programs, withdrawal must be accompanied by a serious commitment, expressed in policy, institutional and budgetary terms to a continued and strengthened American role".

The UNESCO executive board is holding an emergency meeting this week to decide how to adjust its programme in the light of US withdrawal, which took 25 per cent of UNESCO's operating budget with it. State Department officials say that the US withdrawal will have no immediate effect

on the participation of US scientists in UNESCO activities; the UNESCO board, however, with only \$2.7 million from the United States from next October, will probably have to cut back on some of the programmes with which the United States hopes to remain involved.

Congressional action remains a wildcard. The House Foreign Affairs Committee is scheduling hearings for later this month.

Stephen Budiansky

Japanese students

Choices change

Tokyo

WHEN it comes to deciding on careers, Japan's students seem to share fully their government's faith in high technology. This year has seen a marked shift in university entrance preferences towards courses in biotechnology and the new electronic media, and companies with major biotechnology projects are zooming up the popularity charts for university graduates seeking work.

Part one of the university entrance examinations — the nationwide examinations for state universities which precede those given by individual universities — was held last week. Analysis of the major subjects according to number of applicants shows that students are reacting to the emergence of new industries almost as quickly as the Tokyo stock market. Just like the seemingly endless boom for biotechnology companies there, biotechnology-related courses are attracting more and more applicants. As many of these courses are offered in faculties of agriculture, there is a corresponding rise in their popularity, after some years in the doldrums.

Together with biotechnology, courses related to the new electronic media — the kind of services that will be offered by the information network system — are also attracting more students. The traditional "safe" careers of medicine and dentistry are seeing a sharp decline in popularity.

Much the same picture emerges from surveys of which companies students want to work in after they graduate. Although banks and insurance companies and the giant automobile and electronic makers are always among the favourites, the whisky manufacturer Suntory, which has invested very heavily in biotechnology, is now one of the three most popular companies. The high-technology ceramic maker Kyocera, a leader in the race to develop a ceramic engine, is also enjoying a sudden increase in popularity even though, like Suntory, it is not among Japan's 50 biggest companies.

The popularity of biotechnology extends to the government too. Adding together the appropriations in the budgets of all six ministries with biotechnology projects shows that spending in 1985 will reach 33,000 million yen (US \$130 million), more than 6 per cent up from last year.

Alun Anderson

Casualties in sight

The five UNESCO programmes in which the United States wishes to continue to participate are:

- **International Geological Correlation Program (IGCP)**

About 80 countries participate in this project to formulate worldwide correlations of geologic strata. The programme was established largely because of the efforts of American Earth scientists; some 300 US scientists currently participate in IGCP working groups.

- **Intergovernmental Oceanographic Commission (IOC)**

IOC is the major mechanism for international exchange of oceanographic data. Data on subsurface ocean temperatures and salinity obtained by merchant and research ships are collected and shared; IOC also provides support for the oceanic components of the World Climate Research Program.

- **International Hydrological Program (IHP)**

More than 120 countries are involved in IHP, which is aimed at developing rational management of water resources; the programme is at present concerned with arid, semi-arid and humid-tropic regions. US withdrawal from UNESCO will result in loss of representation in the manning of the programme.

- **Man and Biosphere programme**

MAB is a vast international programme now in its twelfth year, whose stated aim is the study of the natural biosphere and man's effect upon it. Withdrawal will cost the United States its leadership positions in the programme.

- **Natural hazards programme**

This programme has provided US Earth scientists with the opportunity to visit areas prone to earthquakes and other natural hazards; it is directly under UNESCO management. □

Biotechnology

US quarantine rules attacked

Washington

US BIOTECHNOLOGY companies are hoping to persuade the US Department of Agriculture (USDA) to bend its notoriously strict quarantine rules that delay for up to a year the importation of cell lines and cell-line products. The companies complain that unless an exemption is granted, the United States could find itself at a competitive disadvantage in world markets. Industry experts claim that within a few years, US companies will be seeking to import thousands of tissue cultures, hybridomas and tissue-culture products every year.

The USDA quarantine applies to any material prepared with animal products that could harbour the foot-and-mouth disease virus or other livestock diseases not already found in the United States. For the biotechnology industry, this typically means products prepared with bovine serum or the enzyme trypsin, which is obtained from cows or pigs. All imports must go through safety testing at USDA's Plum Island quarantine facility before being allowed into the country. Companies must pay for the test, which usually costs around \$2,500.

Bruce Mackler, general counsellor for the Association of Biotechnology Companies, says that the rules are having an immediate effect on US companies such as Charles River and Damon that are hoping to compete with European companies, notably Celltech in Britain, in the contract fermentation business.

According to Dr Laura Peterson of USDA's Animal and Plant Health Inspection Service, about 100 applications for import permits are filed each year. She said that there has been a noticeable trend among the applications from biotechnology companies towards requests for importing large numbers of hybridomas at a time.

Peterson said she did not know of any cases so far in which biotechnology products had been found to carry disease organisms. But one of the nine outbreaks of foot-and-mouth disease that occurred in the United States between 1870 and 1929 was traced to an imported smallpox vaccine. Foot-and-mouth disease is present in France, Germany, Switzerland and other mainland European countries; it has been eliminated in the United Kingdom, Scandinavia and Japan. Foot-and-mouth is a special case under US law: importation of the virus for any purpose is forbidden. Testing involves inoculating a steer at Plum Island followed by serological testing after 4-6 weeks. A number of less common livestock diseases found in Australia and Africa are also of concern to USDA, but these can generally be spotted with simpler tissue-culture tests.

Peterson indicated that USDA was sympathetic to the industry's concern; one pos-

sibility would be to exempt from safety testing products that will be used in clinical applications or as *in vitro* diagnostics, where the risk of spread to livestock is small so long as used material is properly disposed of. The industry is also trying to encourage development of a tissue-culture test for foot-and-mouth, and is proposing to support a postdoctoral worker at Plum Island to assist in this work; Peterson said that USDA was cooperating in this effort.

CERN economics

Particle physics pays (official)

THE European Organisation for Nuclear Research, CERN (near Geneva), can claim to pay back at least 60 per cent of its current £240 million annual budget in the form of "added value" from its contracts with European industry, claims a CERN report just published*, CERN made a similar study, and claim, for the first time ten years ago. The present report brings the story up to date, projects value forward to 1987 and may be of significance in the struggle over the CERN budget in Britain.

Dr Herwig Schmeid, now an economist at the University of Strasbourg, pioneered the methods the report uses — basically a series of interviews with a random sample of CERN contractors, who were asked to quantify any spin-off from CERN work. Lowest values were always taken, CERN claims, and when the benefit was unquantifiable (such as the prestige of a CERN contract when seeking new clients), "no benefit" was recorded. So the method underestimates spin-off, claims CERN personnel manager Dr Norman Blackburne, one of the co-authors of the new report. It also ignores the actual value of the CERN contract itself.

CERN agreed with the companies investigated that they would not be named, but the report includes some examples of the added value, or "economic utility" as CERN describes it, of CERN contracts, such as the shipyard that contracted for some of the heavy steel work required for the big collision detectors on the CERN proton-antiproton collider. The yard increased the precision of its manufacturing and was able to contract successfully for work on North Sea oil rigs, says Blackburne, and it was also able to avoid laying off some of its workers during a slack period at the yard. But the net benefit to the company was considered unquantifiable, and so not counted in the report.

Analysed by sector of industry, electronics, optics and computing gained most utility from CERN. Over 1973-82, 57 companies interviewed gained six times the value of their CERN contracts in new contracts outside CERN. Sixteen steel and welding companies gained more than five

times, and 22 vacuum, cryogenics and superconductivity suppliers three times. Electrical equipment and precision mechanics manufacturers doubled the value of their contracts, the report claims.

Academic research, in which laboratory animals are used, poses a more difficult problem, one compounded by widespread ignorance of or indifference to the regulations by academics. Mackler notes that "casual transfers" of cell lines have been the norm for years among academic researchers, who are accustomed to putting samples "in their pockets" and blithely walking through customs; no company could take the risk of following that example.

Stephen Budiansky

Industry representatives are meeting USDA officials this week, and a decision is expected soon on new procedures.

One company, not named but probably the Norwegian computer company Norsk-Data, considered its prospects transformed by a large contract to supply the control system for the Super Proton Synchrotron. Another company, employing 50 people, was created just to supply CERN with flanges and clamps for connecting vacuum pipes; it developed a "highly suitable alloy", flexible and radiation-resistant, and from being wholly dependent on CERN now exports 70 per cent of its production.

In general, many companies have used CERN to help them cover the costs of development of a risky new product; or, in one case, where a company was manufacturing racks for electronics, CERN requirements provided an ever-improving standard that enabled the company to beat the competition in more mundane settings such as telecommunications and signalling.

But, says Blackburne, there was no single glamorous breakthrough attributable to CERN contracts: rather an accumulation of expertise. (And in one case, the report admits, the value was negative, where the exacting CERN specifications caused a company to restructure towards providing quality levels that other customers did not require.)

The previous pioneering CERN report on economic utility was criticized for lack of rigour, "but this time we have been extremely strict with our sampling process", says Blackburne. The chief conclusion of the report is that one Swiss Franc spent by CERN on high technology between 1972 and 1983 generated or will generate three Swiss Francs of "economic utility" in 1972-87. The overall cost of the organization in 1972-83 was SF 6,945 million (around £2,300 million), and the estimated economic utility SF 4,800 (around £1,600 million), or 60 per cent of costs.

Robert Walgate

*Economic utility arising from CERN contracts (second study), CERN yellow report 84-14.

Paris museum

Discovery for the people

A GIANT abattoir on a 135-acre site in north-west Paris, now under conversion into a permanent spectacle of science and technology at a cost to the French state of some FF 4,500 million (£450 million), is to be launched in the next few days as an independent "commercial and industrial" organization.

"Launched" is the word, for the *Cité des Sciences et de l'Industrie* at la Villette is surrounded by a great moat, which led to headaches over the access of emergency services solved only by driving a road across two bridges and right down the spine of the building. But this was a minor problem for the originator and present director of the *Cité*, physicist Maurice Levy, who was brought back to get the project under way again after the previous director was beaten by conflicts between the scientific planners of the *Cité* and the architects.

Now Levy says the *Cité* will open in March 1986, and he welcomes its new independence which "will give us much more freedom to deal with partners in industry and the banks". Under fire from all quarters for the cost of the grandiose project, Levy has enough money in this year's budget to complete the building and equip three-quarters of it (leaving room for expansion). But there may be some concern over projected ultimate running costs of FF 500-600 million (£50-60 million) annually.

Levy estimates that 30 per cent will be covered by receipts from 5 million visitors a year, but the rest must be covered by the state — or by industry, which might thus gain too strong an influence over the policy of the *Cité*. And there is a general election next year, in which the government seems likely to shift to the right.

Levy insists, however, that the approach to science and technology at the *Cité* will be "critical". But criticism will be "neutral, discussing the pros and cons; science and technology are no longer taken for granted". So the *Cité* is to emphasize the socioeconomic consequences and context of discovery and technology, Levy claims, but he admits that there is little of this to be seen in a vignette of the coming exhibition now on show on site in the small hall *Janus II*, recently visited by President Mitterrand.

Other features that will distinguish the *Cité* from most science and technology museums around the world are:

- that it will deal essentially with the present and future, and not with the past;
- that it will aim to serve the whole public, and not a specialized community;
- and that it will make use of "the enormous revolution in computing and audiovisual communications technology" since the construction of similar museums (such as the Air and Space Museum in Washington).

Levy's vision for the *Cité*, which will be part of a park including a "Great Hall" for exhibitions (built around a vast 19th-century cast-iron cattle-shed) and the Paris Conservatoire de la Musique (which is moving to the site), is that it should "modernize the minds of the people". The material in the *Cité* will not be presented by discipline (as at the existing Paris *Palais de la Découverte*, which Levy sees as more for science specialists than the *Cité*), but according to people's real-life experiences.

The *Cité* should "make people much more aware of the changes that are taking place, of industrial restructuring, preparing people for the next industrial phase", says Levy of his brainchild.

Considering the costs and implications of this restructuring, the investment in the *Cité*, is small, Levy believes. The site is

peripheral to Paris (unlike the *Centre Pompidou*, a similarly expensive and controversial arts centre designed by British architect Richard Rogers, which now attracts eight million visitors a year and is no longer controversial) but "people are willing to move all over Paris now for cultural events" and underground railway and road links are good. Levy hopes that by the end of 1986, after nine months open to visitors and probably facing a new government, the *Cité* will be able to demonstrate as much success as the *Centre Pompidou*.

But so much for Paris. For the rest of France, the *Cité* will host 40 classrooms continuously for provincial schools to visit for two weeks at a time ("like they do for skiing trips"); Levy is also arranging exhibitions and a variety of events with bodies outside the capital. "We shall be a resource" for the whole of France, says Levy. Sooner or later, everybody comes to Paris.

Robert Walgate

French universities

Appointments in doldrums

THE bigger French universities are facing a major problem over an increase in teaching duties imposed upon them last year — because they cannot find the students to fill the hours they should teach. The result is that "it is impossible to ask for new posts", says one geophysicist affected by the problem at the University of Paris VI.

But according to the Ministry of Education, which is in the midst of a thorough reform of the higher education system in France, the problem varies greatly from university to university, and from discipline to discipline.

Also, as universities such as Paris VI reorganize their entry year (the *premier cycle*) and provide more "orientation" teaching for the incoming students (many of whom leave after the first year), their teaching load will be taken up and they will be free again to request new posts; or so the ministry officials hope.

More teaching, however, means less time for research, unless the universities "modulate" their teaching load by distinguishing, effectively, between teaching and research careers in their departments, shifting more teaching to one professor than another. Under last year's new regulations the universities are free to do this, but, says ministry research director for the universities, M. Bernard Descompes, "I must confess I know no university where the distinction is important".

By discipline, the number of free posts varies greatly because the number of students entering those disciplines varies greatly, and because of government policies of expansion in technology-related areas. In computer science, electronics and all the engineering disciplines, "we need many

more teachers" says Descompes.

On the one hand, in certain areas, the number of new posts offered is so large that "it is impossible to be sure that in all cases we are appointing people of quality". Descompes puts the threshold at 5-6 per cent a year of the total number of staff in a discipline. So in computer science, for example, where French universities are expanding at 10 per cent a year, "we are in danger", says Descompes. Electronics and management sciences also come above five per cent.

On the other hand, in the fundamental disciplines, in particular chemistry, physics and "of course" in literary disciplines, the number of appointments throughout France has fallen to just one per cent. There is a "great danger for the future of those disciplines".

Molecular biology has benefited a little from the connection with biotechnology, and is showing 2 per cent growth in posts, says Descompes. But French biology has been dominated for a long time by traditional disciplines and molecular biology was late to take a grip so the need for expansion in this field is particularly great. In traditional biology, expansion is only 1 per cent. But there is not the great dichotomy in biology that there is in the physical sciences between electronics and basic physics.

Nevertheless, the net recruitment in French universities is now 3 per cent a year, compared with only 1 per cent in 1981, so there has been some progress. "But the situation is still critical in more than half the disciplines", Descompes warns. And in the parlous French economic situation a major increase in appointments does not seem to be on the cards.

Robert Walgate

Southern California degrees

Machine-made grades fraud?

Los Angeles

A CASE of possible computer chicanery is under investigation at the University of Southern California (USC) in Los Angeles.

According to university officials, at least 30 students may have paid thousands of dollars to have their grades changed on transcripts stored in the campus computer system. The higher grades may have enabled some students to gain admission to law school, medical school and graduate school. In some instances, the grade changes meant students could have graduated without repeating failed courses.

Now there is another allegation from United States drug enforcement officers that entirely bogus USC degrees have been sold for up to \$25,000 each.

The university is investigating both charges, according to vice-provost Sylvia Manning. "We take this enormously seriously. We have been embezzled, not for money but for grades and perhaps for degrees", she said. "The degree stands for the knowledge and skill that the university is offering its students."

Two investigations are under way simultaneously. The most recent stems from the arrest last September of a former USC student, Merhdad Amini, who arrived in Louisville, Kentucky, with two kilograms of cocaine in his luggage. Mr Amini, a 27-year old Iranian citizen, is under indictment in Louisville for possessing and intending to distribute the drug. He told a *Los Angeles Times* reporter that he was "set up" by someone who claimed to be his friend.

Drug enforcement officials in Louisville said they learned during the course of the drug investigation that phony USC degrees were being sold.

University officials are now painstakingly comparing transcripts kept in the central computer with duplicate transcripts filed in the various university departments; any discrepancies should eventually be found.

If evidence of tampering is found, the information will be turned over to the Los Angeles county district attorney's office, which is looking into the question of bogus degrees.

A second investigation charging that students have bought higher grades has been under way since last October. Thirty students, whose records were changed "in an unauthorized manner", are due to appear before a university conduct review board later this month. Penalties for those found guilty could range from a reprimand to expulsion from the university or criminal charges of "malicious access to a computer".

USC sources who asked for anonymity said the university's probe is focused on a group of foreign students suspected of selling both the grades and the fake

degrees. The group reportedly paid employees in the university's records and registration office for computer passwords. A gang member would then alter the grades on a transcript or alter a completed transcript by electronically erasing the name of the legitimate student and entering the name of a customer who paid for a bogus degree. In some instances, the sources said, university employees may have done the tampering in exchange for cocaine or money. The investigation has not so far named any suspects. **Sandra Blakeslee**

Research councils

Sterling worries

LAST year's Christmas bonus to the British research councils seems to have disappeared before the ink on the cheques has dried, partly because of the decline of the value of sterling on the international exchanges. Three years from now, according to estimates now current, the Science and Engineering Research Council (SERC) will have to find an extra £8 million to help pay for its overseas subscriptions to organizations such as the European Organization for Nuclear Research (CERN) in Geneva as well as for its own installations overseas, notably the telescopes on Hawaii and La Palma. This will almost exactly cancel out the extra funds won last December.

The history of research council wrangling with the British government over the cost of overseas activities is now rich and lengthening. Three years ago, an unwonted increase in the value of sterling against other currencies allowed SERC to distribute the windfall as an extra capital grant, to the chagrin of the Treasury.

Two years ago, with sterling fallen somewhat, the Treasury agreed that there should be a special investigation of the research councils' extra need for cash. The outcome was a settlement that partially compensated the research council for that year's sterling fall.

The fall of sterling in the past six months, which has been more dramatic, has created a novel set of circumstances. At its meeting last week, the Advisory Board for the Research Councils agreed to draw the government's attention to the problem. In the current year, SERC is having to find £2.4 million out of its budget to pay for the extra costs not covered by the earlier review as well as the larger sum occasioned by the fall of sterling.

Other research councils are not nearly as much affected by currency fluctuations, but have discovered special problems of their own. Thus the Agricultural and Food Research Council is now worrying about the threat that research commissions will tail off during the course of the next three years. □

Hungarian computers

East Europe's freewheeling

COMPUTER skills are becoming important in adult education in Hungary. Last month alone, the Society for the Dissemination of Scientific Knowledge (TIT) launched 15 new programming courses just in Budapest, and Hungarian television began a 16-part course in BASIC, with an optional examination and diploma for successful students at the end, and video-cassettes for those unable to watch regularly. Yet only three years ago, the Central Statistical Office estimated that there were altogether just 2,664 computers operating in Hungary, of which only 919 were of a higher capacity. Now, according to the official news agency MTI, there are between 50,000 and 100,000 personal computers, mostly in private hands and some 200 micro-computer clubs. Approximately a million people (a tenth of the population) have by now encountered computer skills.

This enthusiasm for personal computers is partly a result of government policy, which has been urging computerization of the economy for some 10 years. The cost of mainframe hardware and operation was, however, prohibitive (a computer-hour on a mainframe cost as much as two months' salary for a skilled administrator). The arrival of the "personal" computer (of both Western and socialist provenance) virtually coincided with a new policy of decentralization in all branches of the economy, so that enterprise and farm managers found themselves faced with decision-making of precisely the type where a small computer proved invaluable. The ever increasing small-business private sector likewise found that a computer was, if not a necessity, a potent status symbol.

The Ministry of Education and Culture quickly responded to the new challenge. By September 1983, all secondary schools (in theory) had their own computer and computing skills became a compulsory part of the syllabus. (Vacation courses had to be hastily arranged for the teachers assigned to the new duty, who seem frequently to have been outstripped by their pupils.) Expanding the role of computers in higher education is a constant theme at conferences of educationalists.

The Hungarians, however, are not content with using computers only for work. Computer games are one of the most desirable souvenirs that returning tourists can take home from the West. This, indeed, is true in many socialist countries where youth leaders frequently draw attention to the ideological hazards of space invaders and star wars software. The Hungarian authorities, however, prefer to treat the games as a social vice comparable with tobacco or alcohol, imposing a swingeing 60 per cent duty on all such imports. **Vera Rich**

Poland

Higher Education law revised

THE Polish Higher Education Act of 1982, which embodies a watered down version of the reforms of the Solidarity period, is now up for revision. Last month, a "public debate" was launched on proposed amendments to the act. Discussion is somewhat hampered, however, because no provisional draft of the proposed changes has been officially published. But reports in the underground press claim that the Ministry of Science, Higher Education and Technology wishes to deprive the universities of virtually all the autonomy won in 1980-81 and embodied in the 1982 act.

The minister himself, Dr Benon Miskiewicz, in an interview with the official Polish news agency PAP at the beginning of the present academic year, made barely-veiled threats that the law would be revised if the self-government bodies in higher education continued to exercise their autonomy. Within a few weeks, an alleged draft of the proposed changes began to circulate in academic circles, claiming to be a "confidential" document leaked to Solidarity by a source within the Ministry of Science, Higher Education and Technology. Miskiewicz repudiated the document, claiming that it was simply a "provocation" engineered by "anti-socialist forces". Nevertheless, in a meeting with the Main Council for Higher Education, he outlined a scheme of tentative changes which one participant in the meeting described as "the same only less detailed" than the underground draft.

According to the draft, the proposed changes include:

- Replacement of the open elections for university rectors by a choice between two candidates nominated by the minister; the successful candidate would then nominate the various deans who are also, at present, elected.
- Depriving the university senates and faculty councils ("collegial bodies") of various rights which would then be vested in the selected nominees.
- Removing from the collegial bodies the representatives of students, junior academic staff and auxiliary staff (such as librarians), leaving only senior academics and party and military representatives.
- Empowering the minister to suspend or disband any collegial body, and to dismiss university staff and expel students, if he deems their activities "contrary to the interests of society".
- Reorganizing the students' self-government bodies, described as an obstacle to "comprehensive socialist development" on the basis of the party-linked youth organizations, thereby bringing all matters of student welfare and social activity under party control.

Not surprisingly, former Solidarity activists in the universities and members of the pro-Solidarity Independent Students'

Association (NZS) are opposed to the changes. Equally inevitably, the student "self-government committees" want to put up a fight, although the 1982 act forbids any formal links between the self-government committees of different universities and colleges, giving the minister an additional weapon against them.

More significant is the opposition shown by the Main Council for Higher Education. This is also an innovation of the Solidarity period, and consists of one representative of every higher education institution in Poland (except the Catholic University of Lublin). Many of the delegates to the Main Council are, however, Party members, elected because the universities feel that they might prove effective in putting over the university viewpoint to state and Party authorities without unnecessary hassle. The principal objection of the Main Council is that, even if the changes are necessary, they are premature. The existing law, it says, should be allowed to run for at least

another three years, until the expiry of the term of office of the rectors elected under the law. Chopping and changing so quickly can only destabilize the universities.

Individual academics have spoken out strongly in favour of academic autonomy and the present structure of student "self-government committees"; they include Dr Jozef Gierowski, of the Jagiellonian University of Krakow and Dr Grzegorz Bialkowski of Warsaw who was elected in November in place of Dr Klemens Szaniawski, whose election in May was vetoed by Minister Miskiewicz. Some of their statements and press interviews were made before the official launching of the debate and make no direct reference to changing the act. Nevertheless, their tacit opposition to the proposals is clear. So far, the only group reported to have come out in favour of the change is a "young scientist" conference held earlier this month under the auspices of the official "Association of Polish Students". Even they, however, urged the active participation in academic decision-making by students, junior lecturers and auxiliary staff.

Vera Rich

Research animals

Tiger salamanders perpetuated

Washington

AN *ad hoc* international campaign by retinal physiologists seems to have secured at least a temporary reprieve for the world's only source of captive-bred tiger salamanders. Retinal cells from the tiger salamander, *Ambystoma tigrinum*, are widely used in vision research, and have the potential to become "a classic preparation", in the words of one researcher. The problem is that only one man, 71-year-old Carl Lowrance, knows how to rear in quantity the giant specimens (up to 12 inches long) preferred by physiologists, and his breeding ponds have suffered heavy pollution by sewage and animal slaughter effluent from a local meat-packing plant.

Lowrance is an energetic self-taught natural biologist in the nineteenth century tradition living in Tulsa, Oklahoma. Since inventing the sonar fish locator earlier in his career, he has spent the past 15 years perfecting techniques for culturing salamanders, primarily for use as fish bait. Lowrance says 30 variables must be controlled to rear the animals successfully, from pH and oxygen content of the water to presence of predators. He had been producing half a million salamanders a year before the pollution killed 250,000 animals and made his ponds largely unusable. But Lowrance is unwilling now to go into debt to finance a business expansion, and has had difficulty finding a farmer willing to pay a consulting fee to learn his techniques. "They just copy what they see", he says. "They don't really study and so they don't understand."

Retinal physiologists were alerted to the

threat hanging over their favourite research animal by Dr Ed Pugh of the University of Pennsylvania. Besides the remarkable size of Lowrance's specimens, their value is that they come from a relatively homogeneous breeding population and have known histories. The retinal cells are large (see p.529), and, unlike those of other species, can be maintained in culture. Several Papers based on the tiger salamander preparation are published in this issue of *Nature* (pp. 582 and 585).

Eric Schwarz of the University of Chicago has written that research on photo-transduction is "critically dependent" on supplies of tiger salamanders. Armed with this and other letters of support from prominent physiologists throughout the world, Pugh has persuaded the National Science Foundation and the National Institutes of Health to provide \$26,000 and \$21,000 respectively to set up temporary breeding ponds over the next couple of years or so. The tiger salamander is an endangered species in some states.

Lowrance is now helping to transfer salamanders to the new ponds and is grateful for the scientists' conservation efforts: he takes an active interest in the research supported by his animals. He is now negotiating with several fish breeders who seem interested in learning his methods, and may soon be teaching them. In the longer term, there are some hopes among researchers of setting up a foundation to preserve salamanders — but plans are still a long way off, and it is by no means certain that Carl Lowrance's skills will be preserved.

Tim Beardsley

UK environment research

Council to contract on centre

THE prospectus of the Natural Environment Research Council (NERC), the smallest of the publicly supported British research councils, is more a recipe for internal reorganization than an account of what research the council hopes to sponsor in the years ahead. The prospectus, trendily called the research council's "corporate plan", to be published today (14 February) does not depart significantly from the brief account that appeared earlier in the year (see *Nature* 10 January, p.91).

The first drafts of the corporate plan were circulated to the NERC council nearly a year ago, since when the underlying principles have been reaffirmed at several discussions. Most recently, the report was approved with minor amendments at the council meeting at the end of January. Discussions with unions representing scientific employees of NERC have, however, only recently begun.

The starting point for the council's reappraisal of its future is a gloomy forecast of its annual budget for the years ahead. In the present year, the council has a total income of £91.5 million, of which £65 million comes out of the British government's science budget and the remainder from separate government departments in return for commissioned research. NERC now estimates that its share of the science budget will decline by 3 per cent a year for the next decade, and that the value of its research commissions will decline by 10 per cent a year for each of the next two years, but will thereafter be stabilized (in real terms), largely because NERC plans to pay more attention to the marketing of its services. The result will be, the plan suggests, that total income five years from now will have fallen to £78.5 million (at present values).

Cuts

Within this framework, the plan says, NERC intends to spend a greater proportion of its income (21.5 per cent by 1989 rather than the current 12.5 per cent) on the support of university research and to create more flexibility within its general pattern of operations. This leads to the conclusion that there must be a reduction of the numbers of people directly employed on administration and research (3,130 on 1 October last) to 2,230 in the financial year 1989-90. The plan says that these reductions will be accomplished by "enhancing" present arrangements for voluntary retirement, by enforcing the retirement age of 60 and by increasing the proportions of staff employed on short-term contracts.

The plans says that a "key" element in NERC's manpower policy in future will be the recruitment of all new members of staff on 3-5 year contracts, which will provide "a more adequate probationary period", give the council better control of the proportions of scientists trained in different

disciplines and (with luck) assist in the interchange of people between universities and council establishments.

Capital expenditure is also under pressure. Present commitments and plans would have the council spending a little more than at present (£11.2 million in the present year) on capital equipment in each of the three years ahead, nearly a third of that on behalf of the development of the British Antarctic Survey (whose budget within NERC has been protected since the Falklands War). NERC's plan now says that the council is being squeezed by the need for expensive equipment in fields such as image processing, which is "critical to the development of the environmental sciences", and that doubts have now arisen about the council's capacity to complete the second phase of development of the site to which the British Geological Survey has been decanted (at Keyworth in Nottinghamshire) as well as to "replace ageing research vessels". All plans for capital spending, the plan says, are to be reviewed.

Restructuring

The most contentious part of the plan is that concerned with the management of NERC's affairs, which is to be centralized. The plan's proposals redefine the areas of responsibility of the council's present committees without substantially reducing their number. The chief innovation will be the appointment of three senior scientists to NERC's headquarters as "directors of science", with responsibility for Earth sciences, terrestrial (including freshwater) sciences and marine sciences respectively. The intention is that the holders of these posts should be chairmen of committees concerned with the scientific direction of council laboratories as well as with the oversight of recommendations on the support of university research (including awards for graduate students) put forward by the University Affairs Committee (to be renamed "committee E") and its network of subcommittees (which are to be re-organized).

The new committees will include fewer members of NERC as such (four rather than five), fewer outside advisers but more senior NERC employees, who will be nominated by the new directors of science. The plan does not suggest the terms on which these directors will be employed (short-term or permanently).

NERC's prospectus suggests that the post of secretary of NERC will become more important in the future. On the management of the enterprise, the plan describes the future "chain of senior command" from the chairman through the secretary to the directors of science and "thence to the directors of NERC institutes, units and/or programmes". The directors of council institutes are to report "on all management matters" to the

directors of science, who will be responsible for "proper coordination" with colleagues, the "management" of commissioned research at NERC institutes and of "solicited" research at universities and polytechnics, as well as the "oversight and integration" of research support in response to grant applications from academic institutions. The British Antarctic Survey will be excepted from these arrangements, and will report directly to the secretary of NERC.

The objective of these changes, according to NERC's plan, is to reduce the "heavy burden" on council members, to devolve financial responsibility and to integrate senior NERC management into the decision-making process. One novel feature of the system proposed is that institute directors who become members of one of the three main committees will, by the principles of collective responsibility, thereafter be committed to the decisions arrived at — or at least will be less able to work against them.

Research

Two important issues are not settled by the plan; the degrees to which future income from commissioned research is vulnerable or, alternatively, can be made to grow. The new directors of science are intended to function in part as salesmen for NERC's research services. The report notes that the decline of the percentage of commissioned research is part of the cause of its present troubles, but says that it would not wish the proportion to exceed a half for fear of unbalancing the research programme as a whole. The plan suggests that growth in commissioned research is more likely to come from industry and local authorities than from central government departments. There is no explicit statement on the future of the British Geological Survey, NERC's chief earner of commissioned research income.

The plan also says that the council intends to review and if necessary to change the charter under which it operates at present, partly with a view to shedding "peripheral activities". Quite what the council has in mind is not spelled out, although there are some council institutes and research units that could be transferred to other bodies.

NERC's case for increasing support for academic research is qualitative rather than quantitative. The plan lists research opportunities that have arisen in recent years, laments that past attempts to economize by reducing manpower have been frustrated by the increased cost of equipment and notes that the council has a responsibility for ensuring an adequate supply of manpower in fields such as meteorology where its own research is insubstantial.

The plan emphasizes that the increased proportion of the total budget (21.5 per cent five years from now) to be spent on academic research will be net of research support by means of scientific services centrally provided. □

Testing for AIDS

SIR — I would like to clarify several statements in your news article on anti-HTLV-III testing of blood donors (*Nature* 13 December, p.583). I am director of the American Red Cross Blood Services, Northeast Region, which encompasses only Massachusetts and Maine, not "the north-eastern United States" as your article indicates.

Neither this nor other regional blood services are being rushed by Secretary of Health and Human Services Margaret Heckler. It has been said, and it is possible, that there is pressure from high levels of government to expedite the availability of a test for AIDS (acquired immune deficiency syndrome) infectivity. Licensing must come before availability, and this decision lies with the the Food and Drug Administration (FDA), as you correctly pointed out. Any "rushing" would thus be intragovernmental, or between the government and test-kit manufacturer. Regional blood services will implement licensed testing if mandated by FDA. It is my hope, however, that licensing or mandating will not occur until regional blood centres are in a position to implement the test with reasonable assurance that the number of false positive test results will be minimal, and there is a clear understanding of the significance of anti-HTLV-III test positivity. To my knowledge neither issue has yet been resolved. What we really need is a reliable, quick, simple, inexpensive test for AIDS infectivity (not just presumptive evidence of previous exposure to HTLV-III). The safety of blood transfusion thus could be significantly improved.

The tests being developed have the potential for causing undue concern for a number of healthy donors while not removing from the blood supply all the units that are potentially infective for AIDS. Indeed, data have recently been presented at the American Society of Hematology meetings that indicate that apparently healthy persons can carry HTLV-III virus for significant periods without developing the antibody.

PETER L. PAGE

American Red Cross Blood Services,
Northeast Region,
Needham, Massachusetts 02194, USA

Thinking machines

SIR — Further to your recent leading article on the Reith Lectures (*Nature* 29 November, p.397), many people, particularly those working in artificial intelligence (AI), argue that a conscious/intelligent machine could in principle be constructed on more or less present day lines; that is, it would be a Turing-like machine, with, in Professor John Searle's phrase, a "formal program".

Let us imagine such a machine. Since the program is formal, it can be printed out, in some approximation to normal language, on paper. Now it is the essence of programming that anything the machine can do can in principle be understood in terms of the printed program. Any output(s) generated by the machine in response to any input(s) are implicit in the program, in the marks on the paper. But the Turing argument was basically that if the input-output exchange looks conscious/intelligent, then the machine (and therefore the program) must be regarded as conscious/intelligent. This means that whatever consciousness/intelligence resides in the machine must also reside in these marks. But surely not even the most dedicated AI enthusiast would assert this.

While this argument is perhaps not rigorous, it does seem to me to support Professor Searle's thesis that a formally programmed machine cannot be conscious/intelligent.

J. F. CRAWFORD

Affander 967,
CH-5303 Wuerenlingen,
Switzerland

Numerology

SIR — In July 1983 (*Nature* 304, 11; 1983) I presented a series of numerological results concerning the masses of elementary particles. My article was accompanied by a sceptical commentary, in which you gave the view that the results are likely to be no more than coincidence and are in any case a "fruitless pursuit". Whether or not it is a fruitless pursuit is, in my view, still very much open to debate, although I doubt whether any such debate would reveal anything other than a wide variety of subjective opinion on the matter. But the real issue is, or should be, whether or not the results can be dismissed as no more than coincidence. As such, the crucial factor is the statistical evidence.

Among the results in my article were two relating to the combined mass of the particles in the basic meson octet, and the basic baryon octet, relative to the proton mass. The combined mass of the particles π^0 , π^+ , π^- , K^+ , K^- , K^0 , and η , of the basic meson octet, is 3.13935 times the proton mass. The multiplier here is close to π , or 3.1415926. Such a close relationship clearly has a specific chance of having arisen. 3.13935 is π to the power of 0.999376, and a little elementary mathematics shows that only one random rational number in 801.2 will have such accuracy relative to any integral power of π .

The combined mass of the particles ρ , η , Λ , Σ^+ , Σ^0 , Σ^- , Ξ^0 , and Ξ^- , of the basic baryon octet, is 9.8146 times the proton mass. The multiplier here is close to π^2 , or 9.869604. Once again, the relationship had a specific chance of occurring. 9.8146 is π to the power of 1.995118, so only one random rational number in 102.4 will have such accuracy relative to any integral power

of π . This means that the odds are 82,000 to one against any pair of random rational numbers occurring with an overall accuracy, relative to integral powers of π , similar to that of the above two results.

Given the simple nature of the results, it seems clear that coincidence is unlikely (though not impossible). Here we have a phenomenon that deserves serious scientific attention, and not a casual dismissal in terms of coincidence. Furthermore, in view of the absence of any full physical theory of particle mass, there ought to be greater attention given to those observations that might provide useful clues pointing in the right direction. Can science really afford to ignore such results?

PETER STANBURY

Flat 2,
3 Ferndale,
Tunbridge Wells, Kent TN2 3RL, UK

Back to Kekulé

SIR — In his Commentary on the philosophy of science, George Gale¹ mentioned that the dialogue between physicists and philosophers of science had recently been augmented by computer scientists and theoretical biologists.

Even more recently, some practitioners of chemistry have also joined in the dialogue: philosophical aspects of surface chemistry², physical organic chemistry³, bio-organic chemistry^{4,5} and quantum chemistry⁶ have been discussed in the literature.

Kekulé⁷ published some thoughts in 1878, and that address is still worth reading today.

F.M. AKEROYD

Bradford and Ilkley
Community College,
Great Horton Road,
Bradford, West Yorkshire, BD7 1AY,
UK

1. Gale, G. *Nature* 312, 491 (1984).
2. Good, R.J. *Chem Technol.* 10, 100 (1980).
3. Zollinger, H. *Chem. Br.* 16, 257 (1980).
4. Akeroyd, F.M. *J. Chem. Educ.* 61, 697 (1984).
5. Akeroyd, F.M. *J. Chem. Educ.* 61, 434 (1984).
6. Weininger, S.J. *J. Chem. Educ.* 61, 939 (1984).
7. Kekulé, A. *Nature* 18, 210 (1878).

Human embryos

SIR — I support the views expressed by John R. Ling (*Nature* 24 January, p.262). As a biologist, I can see no justification for an arbitrary age below which experimentation on human embryos is legitimate, while above this level it is not. Surely the zygote becomes a potential human being from the moment of fertilization and, if so, it is from that moment entitled to as much respect for, and protection of, its life as is accorded by most societies to a child or adult. In this respect, I agree strongly with the minority of the Warnock committee, and it will be interesting to learn the views of biologists and others among the readership of *Nature*.

J.R. BAKER

4 Belvoir Road,
Cambridge CB4 1JJ, UK

US views of British university research

from Alan R. Katritzky

A university researcher with long experience of British universities, but now working in the United States, reflects on the differences between the two systems and advocates modest reforms.

IN 1962 I was appointed as the founder dean of the School of Chemical Sciences of the then emerging University of East Anglia at Norwich, in the United Kingdom. In September 1980, I left the University of East Anglia to take up an endowed chair in the Department of Chemistry at the University of Florida. Now, some four years later, it is perhaps appropriate to look back over my eighteen years at the University of East Anglia and to describe some of the impressions gained during that period in the light of my more recent experience in the United States.

Anglo-American differences

There are some very important differences between British and American universities that have considerable relevance to the way in which scientific research is organized.

In the United States, most of the money for research in science comes from outside the university directly to the principal investigator and derives from approaches commenced at the initiative of that member of the staff to one of numerous government agencies, foundations, and industrial companies. By contrast, in the United Kingdom a larger portion of the financial support for research comes from the University Grants Committee (UGC), which is supposed to ensure that grant applications are made from "well-found laboratories" (a phrase coined by the old Science Research Council). The number of independent granting agencies is very much smaller, with the publicly-supported research councils predominant.

Whereas there are, of course, differences in the reputation and status of different universities in all countries, the range is far greater in the United States than in Britain. Furthermore, a university's position in the pecking order is far more mobile in the United States; both whole universities and individual departments can go up or down in reputation with considerable rapidity.

It follows from these two considerations alone that when a university scientist in the United States is successful in attracting money and resources, it helps all his colleagues without taking resources away from them. Accordingly, there is a very great motivation to attract good men and women into a department, to keep them, to encourage them, to motivate them to continue to do well. In a British university, if one member of staff does very well, it can mean that in one sense he or she takes a large share of the departmental resources, which

can be the source of resentment from other members of the department.

In the United States, there is no such thing as an automatic salary increment or an age-for-wage scale. Every university decides pretty much independently what it is going to pay its various members of staff. Although a portion of the annual increment is designated as a cost of living supplement, much of it is discretionary. The manner in which these discretionary increments are awarded varies from university to university, but broadly reflects the achievement of individuals. For example, at the University of Florida, every staff member is formally assessed by the departmental



Katritzky — back in Florida.

chairman each year, on the basis of how well he or she has done in the three areas of teaching, research, and administration, as determined by a report required from the staff member outlining his or her activities over the previous year. The final decision regarding the level of the salary increase is made by the dean after discussion with the department chairman.

In the United States, most academic appointments are for nine months and salaries are paid only from September until the end of May. Few teaching positions are available during the three summer months, so that to get paid most faculty members are required to solicit and acquire outside support for their work from which they can 'pay themselves' at a rate equal to their rate from the university during the other nine months. This system acts as a great incentive for members of staff to obtain funds from outside sources.

The means of obtaining tenure in the United States is quite different to that applying in Britain. In the United States, tenure is by no means an easy prize. There is quite a rigid distinction between postdoctoral positions, which are positions in the group of a senior (usually tenured) staff member, and assistant professorships. On

appointment as an assistant professor (which equates to the British lecturer), a new faculty member must work completely independently, and publish independently, if he or she is to become tenured. Decisions on tenure are usually made only at the end of five years, during which an assistant professor must have shown substantial productivity and an ability to make an independent reputation.

American weaknesses

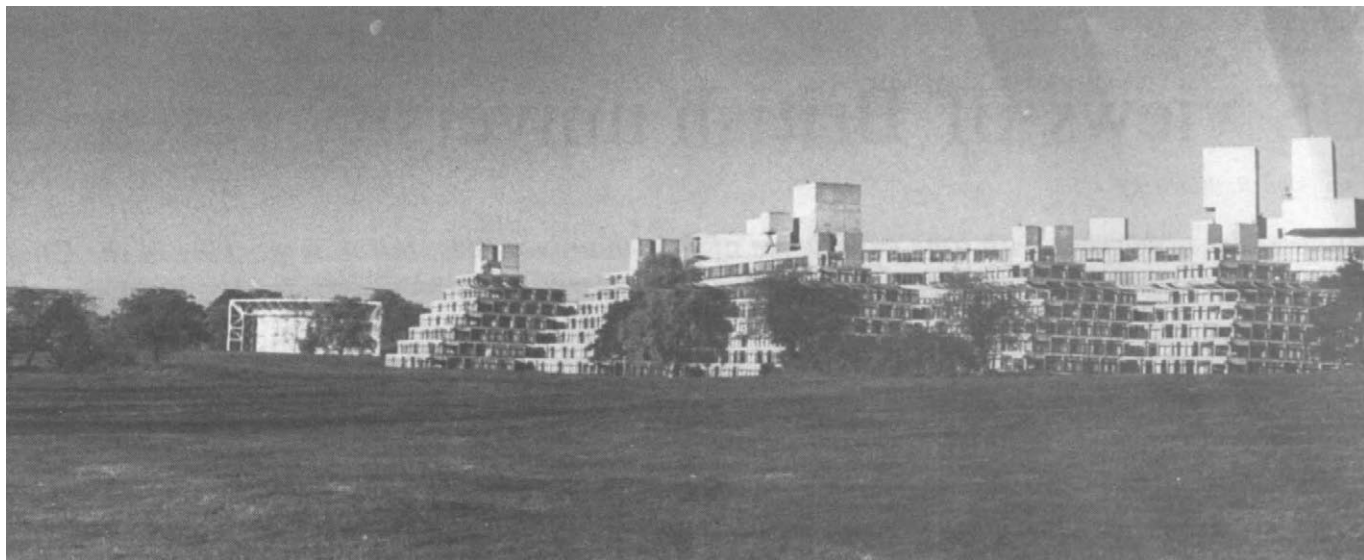
We can learn not only from the strengths of the American system but from its weaknesses, which do exist.

The emphasis on the complete independence of each and every professor (equals lecturer, reader and professor in the British idiom) means that there can be a tendency towards small and nonviable research groups. Even when tenure has been granted, there is an unwritten rule against over-close collaboration between staff members in a single department; for two staff members in the same department to run a joint research group is extremely rare.

One consequence is the lack of continuity within a research group. Graduate students doing research (as distinct from course work) will rarely stay for more than four or five years and postdoctoral fellows for a maximum of about three years. The result is to place a considerable burden on each professor who must provide continuity as group leader, but also, of course, to ensure that he keeps up with the frontiers of knowledge. As soon as he drops behind, his group will tend to disintegrate.

Another consequence of this emphasis on independence is a tendency for some departments to become a collection of small, almost isolated, and sometimes rival, mini-departments, with little cooperation in the use of resources. This may entail the duplication of much equipment and the inefficient use of material and manpower. This consequence of independence is not inevitable; we avoid it at the Department of Chemistry at the University of Florida.

But while the US system tends to give faculty members much more independence within their own departments and their own universities than in the United Kingdom, US researchers are correspondingly more beholden to their peers across the continent as a whole. In other words, the scientific community as a whole exerts much more pressure to conform. So there is a tendency to concentrate research into 'fashionable areas' so as to attract money and to work



The University of East Anglia, Norwich

in fields that will lead to lots of quick publications, sometimes only to find out 'more and more about less and less'. Research on a broad front is difficult to justify and is deprecated in many areas as being 'unoriginal'.

American advantages

Perhaps the overriding advantage of the US system is that American professors have both the motivation and the opportunity to do exceedingly well. They can look towards unlimited horizons, given the ability and the capacity to work hard enough. The acquisition of lots of sizeable research grants brings him personal gain (tangibly, in terms of summer salary), manpower and material resources, acclaim from his colleagues in his own university and from outside—and the opportunity to move to another university on better terms.

Every encouragement is given to such a person to secure outside resources, obtainable from a large number of different sources. The story of the British lecturer who, before a trip to the United States, was told by his department head, "Now, you be sure not to get more money; your group is big enough already", would be quite unthinkable in the United States.

There is also no problem of dead wood in the American system. If an American professor stops publishing, no matter how great his previous accomplishments, he will be cut off from new research resources—dropped cold; certainly a cruel system, but an effective one.

University politics

The scientist interested in research has to give much more of his time to university politics in Britain than in the United States. My own worst memories (among, I should add, many very pleasant ones) of my time at the University of East Anglia are of university committee meetings. These included interminable arguments in a large committee about promotions of colleagues, particularly at the lecturer to senior lecturer

borderline, when only a certain number of posts were open. In the university senate, a tiny minority of self-styled democrats could waste endless time. Worst of all, in my view, were the arguments over finance provoked by continuing crises—the need to call back funds from the operating units, to argue out in detail how these funds should be extracted from budgets to which they had already been allocated and, to a large degree, spent.

In the United States, it appears to be possible much more effectively to shield the research oriented professor from calls on his time for innumerable committees.

What can be done?

It is tempting to recommend that the 'age-for-wage' salary scale for lecturers at British universities should be abolished, together with the system of fixed increments and that there should be introduced a new system of merit increases. But there appears to be little chance of any such far-reaching reform in Britain.

It is perhaps more feasible to suggest that the UGC should change its policy of giving block grants to universities which the universities are theoretically able to use as they wish. The myth of the independence of British universities from government control is a very shaky, while the present position leads in some ways to the worst possible situation. The UGC now gives guidelines which are held to fairly religiously but often only after considerable discussion in university senates. It would surely be possible to set up some method for the rating of individual departments within universities and then of allocating more resources to the best departments and correspondingly less to those departments not doing well. This would provide a powerful incentive to do well and to continue to improve.

It should also be clearly laid down by each university Council that the "democratic" assembly of scholars should decide on such questions as the teaching syllabus,

examination marks and degree classifications, and library acquisitions. It should equally be made clear that democracy does not extend to an equal say in the way that resources are allocated. These resources are provided from tax-payers' money, and the final say in the way that they are allocated is by parliament, which acts through the UGC. The UGC needs to give more precise guidelines in this respect. The trend to equal shares for all is most dangerous.

In those universities where 'democratic election procedures' have come into effect for heads of departments, there should be certain safeguards. If all the faculty are eligible for election, then the wishes of the faculty should certainly be taken very much into consideration, but the final say as to who would be the best department head should lie with the vice-chancellor.

Much of the time presently spent in committees could be saved if the pro-vice-chancellors in British universities were given more responsibility and power to make decisions. For example, in the allocation of financial resources it would seem logical to make an initial allocation between the major areas of the university, say the humanities, the social sciences, the sciences and the technology faculties. Then, within each area, the break-down between the individual departments should be done by a single pro-vice-chancellor in charge of that area, of course, in consultation with the heads of each of those departments, but with the final say in the matter.

The British university system has many excellent aspects, and it is with a view to conserving and improving these that I offer my impressions in a spirit of constructive criticism. □

Alan R. Katritzky is Kenan Professor in the Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA. Professor Katritzky was elected a Fellow of the Royal Society in 1980, while still at the University of East Anglia.

Three-dimensional electron lattice

At low temperatures and in strong magnetic fields, electrons (in a semi-conductor) can indeed be made to crystallize into a lattice, confirming a prediction of Wigner half a century ago.

THE idea that electrons may form into a crystal lattice, originally due to Eugene Wigner, has finally been confirmed after exactly half a century. The delay does not imply indifference. Since C.C. Grimes and G. Adams demonstrated in 1977 that electrons at the surface of liquid helium will form a two-dimensional lattice of the Wigner type, there have been several claims of evidence for three-dimensional electron lattices in materials as different as InSb and the ternary alloy $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$. But only now does it seem to be accepted that T.F. Rosenbaum and S.B. Field of the University of Chicago, working with D.A. Nelson (Honeywell, Lexington) and P.B. Littlewood (AT&T Bell Laboratories) have unambiguously demonstrated a three-dimensional electron lattice in $(\text{Hg,Cd})\text{Te}$ with x at about 0.24 (*Phys. Rev. Lett.* **54**, 241; 1985).

The curious sensation engendered by the notion of an electron lattice is that it seems a contradiction of quantum mechanics. If, as we are always being told, electrons are not localized entities but, at the very least, smudges and perhaps even standing waves, how can they be localized at the vertices of a lattice? The answer is that what matters is only the free energy of the system. The uncertainty principle implies that localization entails a penalty (increased kinetic energy) that may be outweighed by the reduction of repulsion energy stemming from regular arrangement on a lattice.

Naturally, the whole idea would be a more direct assault on more elementary principles were it not taken for granted that an electron lattice can form only if embedded in a neutralizing sea of positive electric charge, such as that provided by a solid crystal lattice from which the electrons have been detached. This is part of the reason why the search for electron lattices has concentrated on semiconductors. The best candidates are materials with only a small concentration of charge carriers of high mobility (or small effective mass compared with that of the electron) which also have a large dielectric constant. Whence $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$.

The other variable that matters is an external magnetic field, the effect of which is to assist the localization of electrons, at least in a plane perpendicular to the field. (Wave packets are elongated in the field direction.) What Rosenbaum and his associates have done is to follow the process of crystallization of electrons into a lattice at low temperatures (between 1 K

and 10 mK) under the influence of an increasing field. Their decisive advantage over earlier workers seems to be the homogeneity of their samples and the low temperatures they have achieved.

Recognizing that the crystallization of electrons is not perhaps as simple as might be expected, but depends on a measurement of the resistivity and Hall resistivity of the sample, the standard means of estimating the mobility and density of charged carriers in a conductor. (Nothing more magical than Faraday's laws of electromagnetism is involved; electric charges moving perpendicular to a magnetic field are also dragged sideways, creating a voltage perpendicular to both the field and the carrier current.)

In these terms, the phenomenon of crystallization is dramatic enough. Rosenbaum *et al.* find that the Hall resistivity of their $(\text{Hg,Cd})\text{Te}$ alloy is only small when the external magnetic field is zero and that it remains virtually constant with increasing field until some critical value of the field (about 6,000 Oe at 10 mK) is reached, after which the resistivity increases rapidly (and linearly) with increasing field.

The critical field at which the supposed crystallization occurs is, as it should be, an increasing (again linear) function of the temperature. The inference, that the sudden loss of mobility is a mark of the onset of electron crystallization is possible largely because of the low temperatures at which the observations have been made, for only then can other explanations be excluded; the trapping of individual charge carriers by donor atoms is the most obvious of these.

For what it is worth, the measurements agree well with calculations, notably those of W.G. Kleppman and R.J. Elliott (*J. Phys. C* **8**, 2729; 1975) who set out by representing localized electrons as gaussian wave packets whose dispersion is different (and greater) in the direction of the magnetic field than in the two transverse directions. The calculations show that circumstances may indeed arise in which the repulsive Coulomb energy of electrons in a regular array exceeds the increase of kinetic energy occasioned by localization. Kleppman and Elliott point out that the physical explanation is really quite simple; the Coulomb energy is proportional to the inverse of the electron separation but the kinetic energy to the square of that inverse, so that the former must dominate the latter

if the separation is large (or the density of carriers low). The calculations show that the exchange energy between a localized array of electrons is negligible (if the density is low enough), so that the exclusion principle is immaterial, which is something of a surprise.

One accompaniment of this argument which in part accounts for present interest in electron lattices is that a crystalline array of electrons should also be spin-polarized, with the electron spins aligned along the magnetic field. Rosenbaum *et al.* have indeed accumulated evidence for such an alignment from the irregularities apparent in the increase of resistivity (not the Hall resistivity) with increasing field strength. Alignment, in their case, occurs at a lower field strength than crystallization. So would it not be possible to demonstrate the crystallization of the electrons directly, perhaps by neutron diffraction? Unfortunately, Kleppman and Elliott concluded some time ago (*J. Phys. C* **8**, 2737; 1975) that the prospects are not bright.

Two sets of issues will now come to the fore. First, the calculation of the phase transition from free to frozen electrons needs attention. So far, people have simply made calculations of the energy of the two phases at zero temperature, so that there is nothing to be said about the dynamics of the transition. The problems clearly, however, have a great deal in common with those arising in the calculation of electron disorder in superconductivity and of the charge-density waves such as are responsible for the electronic behaviour of, for example, polyacetylene and the other materials which are better candidates as organic superconductors.

The practical implications of this development are less easily foreseen. Obviously, it is good that there is now a physical explanation of a phenomenon that may affect the electronic behaviour of an important set of materials. Whether it will be possible to exploit the fact that the lattice electrons are spin polarized in the same direction remains to be seen, but the low temperature at which crystallization unambiguously occurs is an impediment — or a challenge to people's ingenuity. Meanwhile, the demonstration of this phenomenon is yet another proof of the astonishing fertility of Eugene Wigner's contributions, at the outset of his career, to quantum mechanics in particular and to physics in general.

John Maddox

Meteorities

Deep-sea stony spherules and the primordial nebula

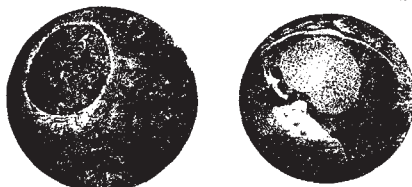
from Derek W.G. Sears

AN EXAMPLE of how work designed to explore one question sometimes throws light on a seemingly unrelated question is provided by a pair of papers just published in *Nature*^{1,2}. The authors of these papers set out to investigate the origin of isolated grains of the mineral, olivine, in the carbonaceous chondrite type of meteorites by making systematic studies of the cathodoluminescence of the grains and their 'minor' element chemistry. The result is both some new thoughts on the origin of the grains and the best evidence yet that some of the stony spherules extracted from deep-sea sediments are extraterrestrial in origin.

Meteorites are relics from the earliest days of the solar system. They are a complicated mixture of components which formed under a variety of physical conditions; some components may even have formed beyond the confines of the solar system. Isolated olivine grains of carbonaceous chondrites have attracted interest because they may be primary crystals formed by condensation in the primordial nebula at relatively high temperatures (above 900 K). Grossman and Olsen³, who were probably the first to make this suggestion, were led to do so by the abundance of refractory trace elements in the grains, the presence of minerals calculated to condense at high temperatures, the absence of glass inclusions and their surface topography. On the other hand, McSween⁴ found that isolated olivine grains have compositional and textural properties which grade into, or are similar to, those of olivine grains found inside another major structural component in chondrites, the chondrules. With a few exceptions, the olivines in chondrules are widely believed to have been formed by crystallization from a liquid. The debate in much that has recently been written on the origin of the isolated olivine grains has therefore centred on whether they are products of high-temperature condensation or the fragments of chondrules—the latter carrying the connotation of crystallization from a liquid.

Steele and his coworkers¹ have examined the cathodoluminescence of isolated olivine grains in two carbonaceous chondrites, Murchison (a type CM carbonaceous chondrite) and Allende (a type CV carbonaceous chondrite). In both cases, the cores of the grains luminesced blue while the rims luminesced red. In general, luminescence colour reflects trace element chemistry, and high quality electron microprobe analyses revealed that the blue regions are low in Fe and enriched in the refractory elements, Ca, Al and Ti, while the red regions are high in Fe and much

lower in refractory elements. The red regions also contain inclusions of glass, silicates and metal, an indication that the olivine formed from a liquid. Steele *et al.* suggest that the blue olivine formed by condensation from vapour at high temperatures and then experienced a short-lived melting event which produced the red rim.



Black spherules, each with a metallic nucleus, collected from *Challenger* at 2,375 fathoms in the South Pacific (left) and 3,150 fathoms in the Atlantic (right). Magnified 60 times. From ref. 7.

This suggestion strikes a familiar chord. In the past few years, several authors have suggested that although chondrules are often droplets produced by melting pre-existing solids, on occasion a few grains of the precursor solid can survive the melting. They are termed relic grains, and are often olivine^{5,6}. Steele *et al.* suggest that all the isolated grains in CM chondrites are relic grains, which were formed by high-temperature condensation and which have subsequently passed through the chondrule-forming process; in essence, both ideas for the origin of isolated olivine grains—vapour condensation and liquid crystallization—are true. The idea merits attention and will suffer intense scrutiny. The immediate problem is that it implies that relic grains in chondrules are commonplace, which is contrary to current observations. Does this reflect our inability unambiguously to recognize all relic grains?

Whatever their origin, the minor element chemistry of isolated olivine grains in carbonaceous chondrites seems to provide them with a unique fingerprint which helps resolve a one-hundred-year-old problem. The circumnavigation of the world by the *Challenger* in the latter part of the nineteenth century marked the beginnings of the modern science of oceanography and yielded many discoveries. One was that the ocean sediment, which was routinely scooped up during the voyages, contains numerous black magnetic spherules, whose discovery was described in 1884 in *Nature*⁷. Their origin remained unclear well into the present century when modern instrumental methods became available. The spherules, which are of the 'iron type', frequently contain an offset iron-nickel core surrounded by magnetite, which is also sometimes nickel-bearing⁸. The prevalence

of nickel is widely considered as evidence for an extraterrestrial origin for the iron type spherules, analogues of the nickel-rich core and magnetite shell structure also being found in iron meteorite fusion crusts. However, the origin of the so-called 'stony-type' spherules remained unclear.

Neutron activation analysis⁹ and defocused-beam electron microprobe analysis¹⁰ of deep-sea stony spherules yield bulk compositions that are consistent with an extraterrestrial origin for the spherules, but interpretation is difficult because of the considerable terrestrial alteration they have suffered. Oxygen isotopes have provided a means of distinguishing the various meteorite classes and finding genetic links between them and, last summer, Mayeda *et al.* produced some oxygen isotope data for deep-sea spherules¹¹. The problem is that not only are the spherules weathered, but they will have picked up considerable atmospheric oxygen during their atmospheric passage. On the assumption that iron spherules contain only terrestrial oxygen and that the stony spherules contain a mixture of terrestrial and extraterrestrial oxygen, Mayeda *et al.* suggested that the extraterrestrial oxygen in the spherules resembles that in anhydrous components from CM chondrites, such as their isolated olivine grains, or CV chondrites.

Like the chondrules, a few of the deep-sea stony spherules contain tiny mineral grains which survived the melting that produced the spherule and so have also been called 'relic'¹⁰. The minor element chemistry of isolated olivine grains from CM chondrites resembles that of the relic olivine grains in deep-sea stony spherules, and is quite distinct from that of olivines from other sources². Take, for example, the Mn-Cr systematics. On a plot of MnO against Cr₂O₃, there is a positive correlation, with a distinctive concave trend, for olivines from the CM chondrites, not found for other classes of chondritic meteorite but rather closely matched by the deep-sea spherule olivines. Unlike the bulk composition and oxygen isotopes of the deep-sea stony spherules, the relic olivines seem to have been impervious to atmospheric or aqueous alteration. This, and the distinctiveness of their composition, provides the best evidence yet that certain deep-sea stony spherules are of extraterrestrial origin. □

1. Steele, I., Smith, J. & Skirius, C. *Nature* **313**, 294 (1985).
2. Steele, I., Smith, J. & Brownlee, D. *Nature* **313**, 297 (1985).
3. Grossman, L. & Olsen, E. *Geochim. cosmochim. Acta* **38**, 178 (1974).
4. McSween, H.Y. *Geochim. cosmochim. Acta* **41**, 411 (1977).
5. Nagahara, H. *Nature* **292**, 135 (1982).
6. Rambaldi, E.R. & Wasson, J.T. *Geochim. cosmochim. Acta* **46**, 929 (1982).
7. Murray, J. & Renard, M.A. *Nature* **29**, 585 (1884).
8. Schmidt, T.K. & Keil, K. *Geochim. cosmochim. Acta* **30**, 471 (1966).
9. Ganapathy, R., Brownlee, D.E. & Hodge, P.W. *Science* **201**, 1119 (1978).
10. Blanchard, M. *et al. Earth planet. Sci. Lett.* **46**, 178 (1980).
11. Mayeda, T.K., Clayton, R.N. & Brownlee, D.E. Paper C3, 47th Annual Meeting Meteoritical Society, Albuquerque, New Mexico (1984).

Derek W.G. Sears is in the Department of Chemistry, University of Arkansas, Fayetteville, Arkansas 72701, USA.

Plant physiology

Ion channels in plant cells

from E.A.C. MacRobbie

AS IN animals, so in plants, the macroscopic electrical behaviour of cell membranes derives from the transient opening and closing of large numbers of ion channels. The technique of patch clamping¹, which involves the electrical analysis of a very small patch of membrane that contains only a few ion channels and is sealed on to the tip of a glass micropipette, allows the characterization of a range of ion channels that are permeable to specific ions, including those channels that are associated with excitability. This method can provide very detailed information on the kinetics of channel gating (opening and closing), and its sensitivity to chemical agents or to the voltage across the membrane. Two recent papers^{2,3} mark the promising extension of patch clamping from animal systems to plant cells.

Patch clamping circumvents the problems that have combined to bedevil electrophysiological measurements in plant cells—the problems of a very tough cell wall, the existence of plasmalemma and tonoplast membranes in series, and the difficulty of gaining access to the cytoplasm. This is particularly a problem in the cells of higher plants where the cytoplasm forms only a very thin layer sandwiched between the plasmalemma and tonoplast with the result that conditions at either membrane are ill-defined and impossible to control. Because of these problems, electrophysiological measurements in cells of higher plants have remained more primitive than those in animal or giant algal cells.

The two papers now published are very different. Moran *et al.*², using wheat leaf protoplasts, show the technique to be feasible, but throw little light on membrane properties in physiological conditions. They present a rather confusing range of behaviour in different membrane patches and solutions, without making it clear whether the type of behaviour observed is a function of the particular patch or of the bathing solution. Out of 41 successful patches, 23 yielded discrete current fluctuations, representing the opening and closing of individual ion channels grouped into four classes according to their conductance (10–20 pS, 35–40 pS, 70–100 pS, and 160–180 pS). None of the bathing solutions used in the illustrated experiments is in the least similar to the normal cytoplasmic (or external) ion composition; rather their choice reflects the importance of sodium channels in many animal cells.

Why do the other 40 per cent of the patches show only smooth currents, with no indication of the opening or closing of individual channels, and with a very high patch conductance? The authors interpret this behaviour, termed “macroscopic”, in

terms of “hot spots” in the membrane, with hundreds of channels, each of conductance less than 5 pS, in the small area of the patch. In the record they show, this behaviour was observed in 25 mM LaCl₃, but it is unclear whether this solution is a prerequisite for the behaviour. The frequency of macroscopic behaviour, if not the result of La treatment, is unexpected. If the conductance of the membrane of wheat leaf protoplasts is in the upper range of values measured in other plant cells, then the existence of even one hot spot would account for the whole cell conductance, and the conductance of a cell with hot spots in 40 per cent of its membrane would be 100–1000 fold higher than that so far observed. Thus, the work on wheat leaf protoplasts signals a major technical achievement, but its significance in a physiological context is not established.

In contrast, Schroeder *et al.*³ start with the firm physiological aim of studying single ion channels in stomatal guard cells (of the bean *Vicia faba*), in which movements of potassium salts are clearly implicated in the turgor changes responsible for opening and closing of the stoma, through which gases are exchanged. The authors identify a specific K⁺-selective channel, with a conductance of 20–27 pS, when the cell membrane is bathed with 30 mM K⁺ on its extracellular side and 228 mM K⁺ on its cytoplasmic side — concentrations which are likely to be reasonably physiological. At a density of one per 15 μm^2 , such channels, open for one-third of the time, could produce potassium fluxes of the observed order. Therefore, they are likely to contribute to the potassium movements associated with guard-cell opening and closing. We eagerly await tests of the sensitivity of this channel's gating kinetics to some of the factors affecting stomatal aperture. Meanwhile, the paper of Schroeder *et al.* is a landmark in the study of stomatal function.

Finally, it is worth considering which of the ion channels that are known to exist in giant algal cell membranes might now be sought in higher plant cell membranes by patch clamping. First, there is evidence for a voltage-sensitive K⁺-channel in the cells of *Chara* and *Hydrodictyon*, which is closed at the very negative membrane potentials generated by the electrogenic proton extrusion pump (the primary system of active ion transport in plant cells) but which opens at low external Ca²⁺, high external K⁺, or at potentials more positive than a threshold value^{4,7}. Second, there is evidence for an H⁺-channel in *Chara*, which is open at high external pH, at membrane potentials close to the proton equilibrium potential⁸; this channel is

involved in the generation of external alkalinity in the pattern of alternating acidic and alkaline zones seen in *Chara* cells taking in bicarbonate ions subsequent to the photosynthetic fixation of CO₂. In addition, two different ion channels are implicated in the action potential in *Chara* cells—a Cl⁻-channel and a Ca²⁺-channel, whose openings allow efflux of Cl⁻ and influx of Ca²⁺ respectively, both contributing to the depolarizing spike but with different gating characteristics^{9,10}. The action potential in *Mimosa*, like that in *Chara*, involves net loss of KCl^{11–13}, and it may well be that channels very similar to those in *Chara* are involved in excitable cells of higher plants.

Thus, there are clearly identified problems awaiting study by the patch clamp technique. The new possibilities for the detailed and well-defined study of cells of higher plants and their range of ion-related processes are very exciting. □

1. Neher, E. & Sakmann, B. *Nature* **260**, 799 (1976).
2. Moran, N., Ehrenstein, G., Iwasa, K., Bare, C. & Mischke, C. *Science* **226**, 835 (1984).
3. Schroeder, J.I., Hedrich, R. & Fernandez, J.M. *Nature* **312**, 361 (1984).
4. Sokolik, A.I. & Yurin, V.M. *Soviet Plant Physiol.* **28**, 206 (1981).
5. Smith, P.T. & Walker, N.A. *J. Membr. Biol.* **60**, 223 (1981).
6. Keifer, D.W. & Lucas, W.J. *Plant Physiol.* **69**, 781 (1982).
7. Findlay, G.P. & Coleman, H.A. *J. Membr. Biol.* **75**, 241 (1983).
8. Bissin, M.A. & Walker, N.A. *J. Membr. Biol.* **56**, 1 (1980).
9. Beilby, M.J. & Coster, H.G.L. *Aust. J. Plant Physiol.* **6**, 323 (1979).
10. Lunevsky, V.Z., Zherelova, O.M., Vostrikov, I.Y. & Berestovsky, G.N. *J. Membr. Biol.* **72**, 43 (1983).
11. Oda, K. *Plant Cell Physiol.* **17**, 1085 (1976).
12. Toriyama, H. *Cytologia* **20**, 367 (1955).
13. Abe, T. *Bot. Mag. (Tokyo)* **94**, 379 (1981).

E.A.C. MacRobbie is in the Department of Botany, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

A rod photoreceptor from the retina of a tiger salamander with the outer segment, which is only 10 μm in diameter, held in a suction pipette. A patch pipette (right) is used to rupture the membrane of the inner segment so that substances such as calcium chelators or cyclic GMP can be introduced into the cell. Results from these and similar experiments (see pages 579–587 of this issue) provide strong evidence that cyclic GMP is the internal messenger for phototransduction rather than calcium, complementing the report published in *Nature* three weeks ago (24 January, p. 310), when the topic was also discussed in these columns (p. 264). (Photo courtesy of T. Lamb).

Prehistoric art

Ice Age drawings on open rock faces in the Pyrenees

from Paul G. Bahn

THE most dramatic find for many years in the field of Ice Age art has recently been made public both in the 'Palaeolithic Art of Mediterranean France' exhibition at Carcassonne¹ and at the 'International Conference on Palaeolithic Cave Art' at Périgueux. It consists of engravings on a large rock, high in the eastern Pyrenees. They were discovered by Jean Abelanet, who had long been exploring the area in search of megalithic monuments and rock engravings of the proto-historic and historic periods; the site and its figures have been studied by Dominique Sacchi, the foremost expert on the early prehistory of the region.

It has been known for decades that the cave art or parietal art of the Ice Age is not confined to caves. Magnificent sculptured friezes have been found in French rock-shelters, such as Angles-sur-l'Anglin and Cap Blanc². Large sculptured or engraved blocks have been found in other rock-shelters or in sites at the foot of cliffs, such as Roc de Sers or Fourneau du Diable. A few years ago, some deep and eroded, linear and zoomorphic engravings were found in cave mouths and shelters along the River Nalón, Spain³, and other examples of exterior deep incisions are known in similar locations in Cantabrian Spain and the Spanish Mediterranean coast. Linear incisions in and around several caves on Mount Carmel, Israel have been attributed to the end of the Pleistocene⁴ and may be far older. And, of course, it has long been known that Ice Age man lived in open-air encampments far more than he did in caves, and many pieces of portable art have been found in such sites in different parts of Europe. But never before has art of the period been found in isolation, in such an open-air, high-altitude site.

Ironically, this unique find has been made in an area which, through lack of exploration, was almost a void on the palaeo-

lithic map until recent years⁵. The investigation of Abelanet and Sacchi have already gone a long way towards filling the void, and the area is now known to contain a number of palaeolithic caves and open-air sites, as well as one decorated cave, Cova Bastera⁵, which has some red dots and lines on its walls. This is still a pale shadow of the multitude of rich sites and abundant art-caves just to the west in Ariège and the Central Pyrenees; but the new discovery at Campôme gives Roussillon (or French Catalonia) special importance.

The engravings are on a huge block of schist (see figure) which lies, unconnected to the bedrock, at an altitude of about

750 m, to the north of the town of Prades, between Mont Canigou and the sea. The rock has been greatly weathered by wind but, because the eastern face is sheltered, its engravings, although eroded, are clearly visible. This face is covered in engravings, drawn in all directions, and comprising about ten animals — none complete — as well as 'signs' and zigzags. The finest figure is the 7.5 cm high head of an isard, or Pyrenean chamois, situated close to the hindquarters of another herbivore (see figure). The isard still lives in the Canigou massif, whereas the ibex, which features in another engraving, disappeared from the area in the eighteenth century AD.

It cannot, of course, be proved that these engravings date from the Ice Age but there is nothing even remotely similar to them among the many schematic engravings found on rocks in the region by Jean Abelanet⁷. Nor do they resemble the Middle Stone Age and later art of the Spanish Levant which extends into Spanish Catalonia. Moreover, had the Campôme

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

The engraved rock at "Fourols Haut", Campôme, Pyrenees-Orientales (below) and an isard head, 7.5 cm high, engraved on it, close to the hindquarters of another herbivore. (Photographs and tracing courtesy Dominique Sacchi).

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

engravings been found in a cave or on a portable slab they would immediately have been accepted as palaeolithic, probably of the Middle Magdalenian of the Pyrenees (12–11,000 BC). Although their unique location arouses some doubt, specialists at the Périgueux conference were unanimously agreed that the engravings are almost certainly late palaeolithic. Study continues, but one can already point to a certain resemblance to a larger head of an isard deep inside the cave of Massat (Ariège)⁸.

In short, just as we had to adapt our view of 'cave-man' as more open-air sites were discovered, and just as 'cave-bears' were by no means restricted to caves (they probably used them primarily for hibernation), in the same way we must adapt our view of 'cave-art'. Although Ice Age ritual has become identified with caves and rockshelters, because these are the places where the evidence has most successfully survived, it has always seemed probable that Ice Age art, mime, song and dance involved outdoor activities⁹. Now, at last, we have

proof that large surfaces far from caves and shelters were also decorated in the same period, and that such art can survive erosion if the conditions are right.

Once the survival of Ice Age art on cave walls was proved at the turn of the century, there was a surge of exploration and remarkable finds. It is to be hoped that the discovery at Campôme will have similar effects. □

1. Sacchi, D. *Exhibition catalogue, L'Art Paléolithique de la France Méditerranéenne*, Musée de Carcassonne, 1984.
2. Leroi-Gourhan, A. *Préhistoire de l'Art Occidental* (Mazaud, Paris, 1965).
3. Fortea Pérez, J. *Zephyrus* 32/33, 5 (1981).
4. Ronen, A. & Barton, G.M. *Quartär* 31/32, 121 (1981).
5. Bahn, P.G. *Pyrenean Prehistory* (Aris & Phillips, Warminster, 1984).
6. Sacchi, D. *106e Congr. Nat. Sociétés Savantes (Perpignan, 1981)* 9 (1984).
7. Abelanet, J. *Travaux Inst. Art Préhist. Toulouse* 10, 5 (1961); 11, 5 (1962).
8. Gailli, R. *Caougnon: Bull. Spéléo-Club du Haut-Sabartez* 14, 18 (1984).
9. Bahn, P.G. *Proc. Preh. Soc.* 44, 125 (1978).

Paul G. Bahn can be contacted through the School of Archaeology and Oriental Studies, University of Liverpool, Liverpool L69 3BX, UK.

Cell proliferation

Interferons and oncogenes

from Mike Clemens

INTEREST has grown in recent years in the ability of the interferons to modulate cell proliferation and differentiation in cultured cells and to slow the growth of tumours in animals and man. Much evidence indicates that the biochemical mechanisms underlying these effects may well be different from those which bring about the antiviral actions of the interferons—although there may be some overlap. Not surprisingly, the growing awareness of the role of cellular proto-oncogenes in regulating both the ability of cells to proliferate and the expression of transformed or malignant phenotypes has raised the question of whether interferon modulates the expression of the proto-oncogenes (such as *c-myc* and *c-Ha-ras*) in target cells.

The past year has seen a burst of activity in this direction and new evidence of how cell proliferation might be controlled by interferons acting as 'negative growth factors'. So far, the work has concentrated mainly on the expression of *c-myc* in Daudi (Burkitt lymphoma) cells and of *c-Ha-ras* in transformed mouse fibroblasts. The concentration of *c-myc* mRNA is diminished in Daudi cells within 24 hours of exposure to human β -interferon¹. Treatment of transformed 3T3 fibroblasts with mouse β -interferon produces revertant colonies in which there is a significant reduction in *c-Ha-ras* mRNA². The expression of a *c-Ha-ras* and *c-src* is similarly reduced in RT4 bladder carcinoma cells³. In the case of *c-Ha-ras* and *c-src*, at least, the levels of the corresponding protein products (p60 and p21 respectively) are also lowered by interferon treatment. In all three examples,

the effects are accompanied by inhibition of cell proliferation.

These observations are of great interest but raise many further questions. Are the changes in proto-oncogene expression a cause or a consequence of the inhibition of cell proliferation? Are they related to the changes in progression through the cell cycle that interferon can sometimes induce? How do they relate to other changes in gene activity and to the effects of interferons on cellular differentiation and expression of the transformed phenotype in tumour cells? It is also crucial that we understand the molecular mechanisms that link the binding of interferon on cell surface receptors to the modulation of proto-oncogene functions in various cell types. Some answers to this formidable array of problems began to emerge at a recent meeting, held in Heidelberg^{*}.

In view of the high sensitivity of Daudi cells to the growth-inhibitory effects of interferons⁴ and the extensive information concerning the reciprocal chromosomal translocations that involve *c-myc* in these cells⁵, it is not surprising that most progress has been made with this system. At the meeting G.J. Jonak (E.I. du Pont de Nemours, Wilmington) extended his published data to show that the concentration of the 2.4 kilobase *c-myc* mRNA falls by as much as 60 per cent after only 3 hours of treating Daudi cells with interferon. He suggests that this is due to decreased processing or decreased stability of *c-myc* transcripts, because he finds no

evidence of decreased *c-myc* transcription. Such an effect on *c-myc* mRNA stability is fascinating since, in Daudi cells, this RNA already turns over with an unusually short half-life (B. Lebleu, University of Montpellier). Data reported by A. Kimchi (Weizmann Institute), however, do suggest inhibition of transcription of *c-myc* in interferon-treated Daudi cells (see also page 597 of this issue of *Nature*).

The results from these laboratories (and from J. Kelly, Imperial Cancer Research Fund, London) suggest that the changes in expression of *c-myc* are sufficiently rapid not to be a result of the partial accumulation of Daudi cells in the G₀/G₁ phase of the cell cycle that occurs as a relatively late response to treatment with β - or α -interferon^{6,7}. Since there is a relationship between *c-myc* expression and the position of cells in the cell cycle in other systems⁸ it is tempting to speculate that the *c-myc* gene product (p62—a DNA binding protein) is required for some aspect of DNA replication. Indeed, DNA replication is disrupted by interferon treatment of Daudi cells in a number of surprising ways, including inhibition of ligation of Okazaki fragments and decreased stability of newly replicated sequences⁹. But until the level and activity of the p62 protein (rather than of its mRNA) has been measured as a function of time of interferon treatment, its possible role in these effects cannot be judged.

In spite of these results, it seems unlikely that changes in *c-myc* expression are the primary (or the sole) cause of the inhibition of cellular growth by interferons. Kimchi reported (and see page 597) that interferon treatment of (albeit, less sensitive) cells such as U937 (histiocytic lymphoma), HL-60 (promyelocytic leukaemia) or Friend erythroleukaemia cells inhibits their proliferation without any change in *c-myc* mRNA (or in cell cycle distribution). Furthermore, P. Lengyel (Yale University) reported that, although interferon blocks the stimulation of proliferation that follows the treatment of Balb C/3T3 cells with platelet-derived growth factor, it does not block the induction of *c-myc* expression⁸ that precedes the proliferative response. These results serve to emphasize the multifactorial nature of cell growth regulation and malignant transformation, as well as the multiplicity of biochemical actions of which the interferons are capable—a problem very familiar to those who are attempting to unravel the antiviral mechanisms activated by the interferons.

As stated earlier, there is clear evidence that interferon treatment can sometimes inhibit the expression of the human *c-Ha-ras* gene. At the meeting R. Friedman (Uniformed Services University, Bethesda) reported further that transformation of 3T3 cells, transfected with T24 human bladder carcinoma DNA (which contains an activated *c-Ha-ras* gene¹⁰) or with the viral *v-Ha-ras* or *v-mos* genes, is inhibited by interferon treatment. However, the

* Third TNO/ISIR Congress on the biology of the interferon system, Heidelberg, FRG, 21–25 October 1984.

phenotype of an established T24 transformed line, RS504, is not reversed by interferon. It is not yet clear whether these effects reflect regulation by interferons of proto-oncogene transcription or of mRNA stability. If the former, then interferons may be capable of highly selective repression of transcription, controlled, for example, by only certain combinations of promoter and enhancer sequences; indeed the synthesis of other mRNA species such as that coding for 2',5'-oligoadenylate synthetase is strongly stimulated by interferon treatment¹¹. On the other hand, selective turnover of mRNAs would imply a differential sensitivity to nuclease(s) that may be activated or induced by interferons. Perhaps the 2',5'-oligoadenylate-RNase system, known to be important in the

inhibition of replication of a number of viruses by interferons¹², may yet have a subtle role in the anti-cellular actions of these negative growth factors. □

1. Jonak, G.J. & Knight, E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1747 (1984).
2. Samid, D. *et al. Biochem. biophys. Res. Commun.* **119**, 21 (1984).
3. Soslaw, G. *et al. Biochem. biophys. Res. Commun.* **119**, 941 (1984).
4. Adams, A. *et al. J. gen. Virol.* **28**, 207 (1975).
5. Dalla-Favera, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7824 (1982).
6. Silverman, R.H. *et al. Eur. J. Biochem.* **126**, 333 (1982).
7. Gewert, D.R. *et al. Eur. J. Biochem.* **139**, 619 (1984).
8. Kelly, K. *et al. Cell* **35**, 603 (1983).
9. Moore, G. *et al. Eur. J. Biochem.* **139**, 27 (1984).
10. Capon, D.J. *et al. Nature* **302**, 33 (1983).
11. Shulman, L. & Revel, M. *Nature* **288**, 98 (1980).
12. Lengyel, P. *Ann. Rev. Biochem.* **51**, 251 (1982).

Mike Clemens is in the Department of Biochemistry, St. George's Hospital Medical School, London SW17 0RE, UK.

Palaeoanthropology

Primate and human phylogeny

from Eric Delson

IT WAS heartening for an advocate of the merits of cladistic analysis, to see at a recent conference* just how widely it is now being used in the study of primate and human evolution. The purpose of the meeting was to provide a review of the main points of contention and consensus, often pairing proponents of both sides of an argument, an arrangement that led to lively debate on theoretical aspects of evolutionary tempo and mode as well as on the details of presentations.

One focus was on the evolution of the anthropoid or 'higher' primates. L. Aiello (University College, London) reviewed the evidence for a close phyletic relationship between tarsiers and anthropoids, which has been strengthened by work on several body systems in recent years. She supported the concept of a monophyletic (single-origin) Haplorhini for these groups, as did A.L. Rosenberger (University of Illinois, Chicago). He proceeded to the question of whether all monkeys and apes are descended from a single tarsier-like proto-anthropoid or whether the New World platyrrhines have a pre-anthropoid ancestor separate from the Old World catarrhines. By demonstrating the homologous nature of novel (derived) cranial structures in all anthropoids, he provided clear support for the concept that anthropoids represent a clade, not a poorly associated grade of primate evolution. In one of the few new concepts presented at the meeting, Rosenberger offered a unified model for the evolution of anthropoid postorbital closure and facial/neurocranial attachment: increasing masticatory forces in an early haplorhine with a narrow interorbital septum required buttressing, which was provided by the complete postorbital sep-

tum and fusion of the frontal bone and mandibular symphysis.

Several of the best papers came from researchers who reviewed areas outside their own narrow expertise. Thus, palaeontologist P. Andrews (British Museum, Natural History) summarized the molecular biological evidence bearing on hominoid (ape and human) relationships, concluding that the majority of approaches have been essentially phenetic. Immunological and DNA hybridization studies reveal similarities among taxa studied, but whether they are derived or even homologous similarities is unclear; studies of genome sequence data appear more promising. A. Friday (University of Cambridge) concluded from new statistical analyses of mitochondrial DNA sequences in living hominoids that chimpanzees may have had an ancestor in common with humans slightly more recently than either did with gorillas; however, the alternative tree, which has chimpanzees closest to gorillas, was hardly less parsimonious. In discussion, several workers played down the likelihood of a simplistic, linear 'molecular clock' for dating the divergence points identified from the many approaches now employed.

Dental evidence figured strongly in several reports. L. Martin (University College, London) summarized his studies (*Nature*, in the press) of enamel growth and ultrastructure in hominoids. Thick molar enamel caps probably characterized the common ancestor of the great apes and humans, but thickness was reduced in the lineage leading to chimpanzees and gorillas. The rate of enamel deposition slowed in the ancestor of these animals and in the orangutan clade which split from the chimp/gorilla/human line some 15 million years ago. There is some argument whether to recognize each of these clades as a full family or as a subfamily of Hominidae,

with support fading for the previous classification of a strictly human lineage in Hominidae and all apes in Pongidae. In a timely reminder that detailed morphological analysis is required to test evolutionary hypotheses, P. Butler (Royal Holloway College, London) suggested examination of dentine morphology to provide a partially independent source of data and opined that the ancestry of the enigmatic Italian Miocene catarrhine *Oreopithecus* might be sought in relatives of the Early Miocene East African *Rangwapithecus*.

The relationships of *Australopithecus* species to each other and to the origin of *Homo* were at the centre of debate in sessions on human evolution. F.E. Grine (State University of New York, Stony Brook) finds that the deciduous dentition of australopiths reveals a 'robust' clade within which the Kromdraai type sample of *A. robustus* can be separated from the more derived *A. crassidens* and *A. boisei*. M.C. Dean (University College, London) reported a number of derived features shared by *Homo* and robust australopiths (*Paranthropus*) not seen in apes or *A. africanus* (and *A. afarensis* where known): early eruption of relatively small incisors and similarities of cranial base flexion and temporal rotation are the most important. B. A. Wood and A. T. Chamberlain (Middlesex Hospital Medical School) could not easily determine the relationships of *A. africanus* in a 'cladistic' morphometric analysis of australopith and *Homo* crania, but *A. afarensis* was placed closer to the robust species. These results and those of Olson (in *Ancestors: The Hard Evidence*, ed. E. Delson, Liss, in the press) suggest to me that the mosaic pattern of early human cranial evolution is not yet fully deciphered: some features of *A. afarensis* presage the robust australopiths, as do some of *A. africanus* (see Rak, Y., *The Australopithecine Face*, Academic, 1983), but in other ways the latter species shares derived features with *Homo*. Perhaps Wood and Chamberlain's more parsimonious tree, with *afarensis* linked to the robust species and *africanus* to *Homo*, is less absurd than it might seem at first glance.

C.B. Stringer (British Museum, Natural History), who usually studies fossil varieties of *Homo sapiens*, bravely trained his sights farther back in time to provide support for the previously under-appreciated view that more than one species might be hiding among specimens allocated to *H. habilis*. He reported the greater similarity to later *Homo* of the supposedly-female, small-brained specimens (ER 1813, OH13) than of the larger-brained 'males' (ER 1470), especially in characters usually associated with male individuals. Thus, increasing recognition that great sexual dimorphism is to be expected in earlier human groups must be tempered with an appreciation of inter-individual and inter-regional (as well as temporal) differences in evaluating the observed variation and discerning evolutionary relationships and patterns. Only by

*Major Topics in Primate and Human Evolution', sponsored by the Anatomical Society of Great Britain and Ireland and the Primate Society of Great Britain, London, 17-18 December 1984.

assessing the evidence for each source of variation in combination with that for homologous shared-derived features can the complexity of human lineage splitting and adaptations be resolved.

The diversity of viewpoints presented at this meeting and at several similar gatherings over the past year, as well as the recovery of new fossils of early apes and humans (not to mention skeletons of *Homo erectus*) in Africa and Asia during the same interval, testify to the renewed vigor of modern palaeoanthropology. Just 60 years ago, *Nature* published Raymond Dart's account of *Australopithecus africanus*; to bend a phrase, we've come a long way from the Taung baby. □

Eric Delson is at the American Museum of Natural History and Lehman College, City University of New York, New York 10468, USA.



100 years ago

A SCIENTIFIC VIEW OF THE COAL QUESTION

It is well known that our stock of coal is not an infinite quantity, and cannot last an indefinite period of time. It has not been by the mere existence of large quantities of coal in this country, nor entirely by the sale of coal to foreign nations, that so much of our wealth has been obtained, but largely by the circumstance that we were the first nation to apply coal to industrial purposes on a large scale and in a great variety of ways. Other nations also possessing coal, perceiving the great success of this method, followed our example, have overtaken us, and have now rendered it increasingly difficult, year by year, for us to maintain our position as manufacturers.

It would be wise, therefore, boldly to face this serious prospect and consider by what means our national prosperity can be maintained as our coal diminishes in quantity and increases in price, especially as our population is continually increasing, and require to purchase greater supplies of foreign food.

There does exist another and inexhaustible source of wealth and progress, viz. new knowledge obtainable by means of scientific research. It is upon knowledge, and substances, that we must sooner or later depend as a fundamental source of national prosperity.

The practical value of new scientific knowledge as a source of wealth and progress is incomparably greater than that of all the coal-deposits, petroleum springs and gold-fields of the earth. This great truth, though familiar to scientific investigators, is but little perceived or appreciated by our rulers, or by the mass of their electors; and the chief reason for this is the fact that they possess in sufficient knowledge of science. Even Governments can only appreciate that which they understand, and can only act as circumstances and public opinion allow them, and when fettered by an ignorant population, are powerless to preserve a nation from decay. From *Nature* 31 357, 19 February, 1885.

Sedimentology

Hummocky sand deposits generated by storms at sea

from I.N. McCave

AMID the successes sedimentologists have had in providing criteria for the recognition of ancient sedimentary environments, the shallow marine environment has been slow to respond to sustained study. Nevertheless, two broad classes of shallow marine deposits are now recognized according to whether they are dominated by tidal currents or by storm-generated waves and currents. Tidal-current deposits are distinguished by several features, including thick sets of cross-bedding, systematic variation in bedding thickness and evidence of reversing currents in 'herring-bone' cross-bedding. Storms are thought to produce 'hummocky cross-stratification' (HCS), a structure described and recognized for its significance by Harms *et al.*¹ in 1975. Although HCS is thought to be a storm-generated structure, no one has ever seen it being formed and thus there is some debate about its conditions of origin. That debate is further enlivened with the analysis presented by P. Allen on page 562 of this issue².

The structure consists of curved laminations forming hummocks and intervening swales, with low-angle intersections of less than 15°. Hummocks are spaced by 1–6 m, are 0.1–0.5 m in height and usually lack preferred orientations or asymmetry.

Allen asks whether HCS can be formed by waves alone and concludes that the spacing of hummocks is much greater than that expected for low-steepness symmetrical bedforms using plausible wave-orbital diameters. Thus, some other component of flow is required to form HCS; waves may be necessary but they are not sufficient. His opinion is at odds with some contributors to a symposium on sedimentology of shelf sands held in Calgary, last June. In particular, Southard³ and Duke with Leckie⁴, concluded that the structure can be accounted for by waves alone. (It is acknowledged that a unidirectional component is likely to be present as well with wind-generated waves.) Thus there is a difference of opinion which will need to be evaluated when all the work is published.

In any case, the necessity of waves for the formation of the structure limits its occurrence. The necessary bottom orbital velocities occur at 50–100 m depth, at most, and then only on exposed continental margins such as the modern shelves off the Hebrides or the north-west of the United States⁵. In situations with limited fetch, such as the Cretaceous Western Interior seaway, HCS are unlikely to have formed at depths greater than 50 m.

In this light, it is intriguing that sand layers showing HCS are interbedded with

layers that appear to be turbidites, (deposited by turbidity currents)^{6,7}. This has raised a hornets' nest of controversy. To some observers, the beds are not genuine turbidites but have simply been produced by a day or two of strong wind-driven shelf flow and have obtained their grading during the waning phase. There is some support for this view in the presence of graded units off the Yukon Delta, which are supposed by Nelson to be of storm origin⁹. The opposing view is that they, and the ancient graded beds interbedded with HCS, are genuine turbidites¹⁰. How can a turbidity current with a velocity of up to 10 metres per second be generated with a head of only 50 m when Bagnold has estimated that at least 700 m are needed⁸? Will all turbidite formations have to be examined for possible signs that they originated in shallow water.

Any controversy like this contains a rich potential for new insights and interpretations. Working out the dynamical significance of hummocky cross-stratifications is just the beginning. The initiation and dynamics of turbidity currents and of storm-driven flows in ancient shelf seas are the next candidates that will have to be (re)evaluated. □

1. Harms, J.C., Southard, J.B., Spearing, D.R. & Walker, R.G. *Soc. Econ. Pal. Min. Short Course notes*, no. 2, (1975).
2. Allen, P.A. *Nature* 313, 562 (1985).
3. Southard, J.B. in Harms, J.C., Southard, J.B. & Walker, R.G. *Soc. Econ. Pal. Min. Short Course notes*, No. 9 (1982); *Abst. Can. Soc. petrol. geol. Res. Symp., Calgary* 65 (1984).
4. Duke, W.L. & Leckie, D.A. *Abst. Can. Soc. petrol. geol. Res. Symp., Calgary* 32 (1984).
5. Dott, R. & Bourgeois, J. *Bull. geol. Soc. Am.* 93, 663 (1982).
6. Hamblin, D. & Walker, R. *Can. J. Earth Sci.* 26, 1673 (1979).
7. Higgs, R. *Abst. Can. Soc. petrol. geol. Res. Symp., Calgary* 42 (1984).
8. Bagnold, R.A. *Proc. R. Soc. A* 265, 315 (1982).
9. Nelson, C.H. *J. Sed. Petrology* 52, 537 (1982).
10. Walker, R.G. in *Facies Models* 2nd edn (ed. Walker, R.G.) 141 (Geoscience Canada Reprint Series 2, 1984).

I.N. McCave has just moved from the University of East Anglia to the Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.

Errata

In the article by Peter A. Lawrence "Compartment genes in hand" (*Nature* 313, 268; 24 January 1985) references 23 and 24 were omitted. They are:

23. Lawrence, P.A. *Cell* 29, 493 (1982).
24. Lawrence, P.A. & Johnston, P. *EMBO J.* 3, 2839 (1984).

In the article by Jennifer Altman "New visions in photo reception" (*Nature* 313, 264; 24 January 1985) the reference to Miller, W.H. and Nicol, G.D., *Nature* 280, 64 (1979) was omitted.

Intergalactic thermonuclear γ -ray line

SIR—The X-ray emission from clusters of galaxies demonstrates^{1,2} that perhaps as much as $10^{14} M_{\odot}$ of hot $T \approx 10^8$ K gas may often surround galaxies in clusters with a density of perhaps $n = 4 \times 10^{-3} \text{ cm}^{-3}$. If the ion temperature is 10^8 K, the thermonuclear reaction $p + d \rightarrow {}^3\text{He} + \gamma$ should emit γ rays at a rate of roughly $4 \times 10^{41} \text{ s}^{-1}$ with energy $E_{\gamma} = 5.516 \pm 0.016 \text{ MeV}$. Such a source in the Virgo cluster at 15.7 Mpc distance would present a line flux $F_{\gamma} = 1 \times 10^{-11} \text{ cm}^{-2} \text{ s}^{-1}$ at Earth. Although this flux is 10^{-6} of the line fluxes that can be detected by the satellite Gamma Ray Observatory the concept of thermonuclear reactions in dilute space is sufficiently fascinating to warrant a brief discussion.

The total amount of hot gas around a galaxy cluster is still not precisely known; nor is its temperature distribution. Imagine for simplicity an isothermal gas having $T = 10^8$ K amounting to $10^{14} M_{\odot}$ of hydrogen of density $n = 4 \times 10^{-3} \text{ cm}^{-3}$. Coulomb collisions should easily maintain equality between electron and ion temperatures. Prior to astration³, the cosmic deuterium appears to have existed with an abundance ratio $D/H \approx 1 \times 10^{-4}$. The reaction $p + d \rightarrow {}^3\text{He} + \gamma$ will be the most active thermonuclear reaction in such a plasma. The lifetime $\tau_p(D)$ of a deuteron against interactions with protons can be shown⁴ to be

$$\tau_p(D) = [n_p(\sigma v)]^{-1} = 2.8 \times 10^{21} \tau^{-2} e^{\tau} / n_p (\text{cm}^{-3}) \text{ s} \quad (1)$$

where the conventional thermonuclear parameter τ is⁴ $\tau = 8.018 T_8^{-1/3}$ with the ion temperature $T_8 = 10^{-8} T$. If, for example, $T_8 = 1$ and $n_p = 4 \times 10^{-3}$ this lifetime is $\tau_p(D) = 3.3 \times 10^{25} \text{ s}$. Because this time greatly exceeds the age of the Universe there is no question of significant destruction of D by this situation.

On the other hand, the rate of γ -ray emission is $Q_{\gamma} = N_D/\tau_p(D)$, which can be rather large if the number of deuterons N_D is $10^{-4} N_p$ and if the hot plasma is as massive as $10^{14} M_{\odot}$ of protons. For then $N_D = 1.2 \times 10^{67}$ and $Q_{\gamma} = 3.6 \times 10^{41} \text{ s}^{-1}$. This emission rate Q_{γ} is proportional to $n_p/4 \times 10^{-3}$, to $M_H/10^{14} M_{\odot}$, and has a temperature dependence inversely proportional to that of equation (1). It is the dependence of the emission on the properties of the hot intergalactic gas that makes it interesting. Also of interest is the energy of the line feature, which occurs near the sum of the mass release $\Delta mc^2 = 5.493 \text{ MeV}$ and the most effective thermal⁴ energy $E_0 = 23 \text{ keV}$ in the ion plasma; that is $E_{\gamma} = 5.516 \text{ MeV}$. The width of the line feature is somewhat greater than thermal, being⁴ $\Delta E_0 = 32 \text{ keV FWHM}$ at $T_8 = 1$.

The flux is likely to be too small to measure, however. Imagining that hypothetical cloud at the distance 15.7 Mpc to the Virgo cluster leads to a flux F_{γ}

(Virgo) $= Q_{\gamma}/4\pi R^2 = 1.2 \times 10^{-11} \text{ cm}^{-2} \text{ s}^{-1}$. This is 10^{-6} of the line fluxes anticipated measurable by the Gamma Ray Observatory, for example, and it does not seem likely that a permanent low-level observatory in space can ever achieve such a low background level. Nonetheless it is interesting that the order of magnitude $Q_{\gamma} \approx 10^{42} \text{ s}^{-1}$ is of the same order as the Q_{γ} (${}^{26}\text{Al}$) from radioactive ${}^{26}\text{Al}$ in the disk of our Galaxy⁵ that was recently detected⁶ by HEAO 3. The Virgo cluster is unfortunately too distant for us to detect, as are therefore other clusters. The contribution to the omnidirectional universal background can be seen to be well beneath the observed continuum following the line of argument developed by Clayton and Silk⁷. Although this calculation therefore seems at first glance academic, it may be unwise to prejudge the scientific future of the unique ideas involved.

This research was supported by NASA DPR S-10987C for the OSSE spectrometer for Gamma Ray Observatory as funded by subcontract N00014-83-C-2162 from Naval Research Laboratory.

DONALD D. CLAYTON

Department of Space Physics
and Astronomy,
Rice University,
Houston, Texas 77251, USA

1. Fabian, A. C., Nielsen, P. E. J. & Canizares, C. R. *Nature* **310**, 733–740 (1984).
2. Jones, C. & Forman, W. *Astrophys. J.* **276**, 38–55 (1984).
3. Clayton, D. D. *Astrophys. J.* (in the press).
4. Clayton, D. D. *Principles of Stellar Evolution and Nucleosynthesis*, Ch. 4, 5 (University of Chicago Press, 1983).
5. Clayton, D. D. *Astrophys. J.* **280**, 144–149 (1984).
6. Mahoney, W. A., Ling, J. C., Jacobson, A. S. & Lingenfelter, R. E. *Astrophys. J.* **262**, 742–748 (1982).
7. Clayton, D. D. & Silk, J. *Astrophys. J. Lett.* **158**, L43–L48 (1969).

The transmission of AIDS

SIR — Shearer and Rabson¹ have remarked that anatomical differences between the vagina and the rectum may in part be responsible for the frequency with which acquired immune deficiency syndrome (AIDS) is acquired by anal-receptive sexual intercourse; I wish to draw attention to some additional and possibly relevant differences between these organs.

The mucosa of the healthy vagina is much less densely populated by lymphocytes than that of the gut. Only occasional lymph follicles, and no large aggregates such as Waldeyer's ring or Peyer's patches, are found. In addition, to my knowledge there have not been described in the vagina specialized structures analogous to the intestinal M cells — membranous epithelial cells overlying Peyer's patches which take up from the lumen, and present to underlying lymphoid tissue, a variety of antigenic materials including macromolecules and viruses^{2,3}. These contrasts suggest that direct encounters between lymphocytotropic viral agents, specifically

HTLV-III / LAV, and their potential target cells, and thus infection and viral replication may be facilitated in the gut but may occur only rarely, if at all, following simple deposition of virus in the vagina.

With regard to a possible role of semen or sperm antigens in immune dysregulation, it is well known that the intestinal tract is highly active in an extensive system of mucosal immunity⁴. But little has been published on mucosal immunity in the female genital tract. Direct comparisons of antigen uptake, recognition and processing between the gut and the vagina do not appear to have been made. Limited inferences from the literature lead me to suppose that immune surveillance mechanisms may operate at a much more restricted level in the vagina. The hypothesis of semen-induced immune dysregulation^{5,6} should be evaluated in light of existing information and new data on reproductive/genital and lower-intestinal mucosal immunology.

GERALD A. AHRONHEIM

Ste. Justine's Hospital and the
University of Montreal,
Montreal, Quebec H3T 1C5, Canada

1. Shearer, G.M. & Rabson, A.S. *Nature* **308**, 230 (1984).
2. Owen, R.L. *Gastroenterology* **72**, 440 (1977).
3. Wolf, J.L. *et al. Science* **212**, 471 (1981).
4. Bienenstock, F. & Befus, A.D. *Immunology* **41**, 249 (1980).
5. Mavligit, G.M. *et al. J. Am. Med. Ass.* **251**, 237 (1984).
6. Sonnabend, J., Witkin, A.S., & Partilo, D.T. *J. Am. Med. Ass.* **249**, 2370 (1983).

Bird territory and weight gain

SIR — In challenging much of the work by Carpenter *et al* on weight gain of birds in relation to their territory size, Butlin² invokes the possibility that the rate of weight gain is related to the bird's body weight, that is $M = kM$, where M is the weight of the bird and k is a constant. This translates to $\Delta M/M = k \Delta t$, which integrates simply to $\ln M = kt + C$ where t is time and C is a constant. Rearranging this expression gives $M = e^{kt+C}$.

With the realistic assumption that $M = 0$ at $t = 0$, that is at or before conception, we have $C = -\infty$, the modulus of which is substantially greater than kt except after an infinite time. For all practical timespans this results in $M = 0$.

The consequence of this is that the bird has zero weight until an infinite time has passed whereupon it grows at an exponential rate — although by that time it probably wouldn't notice the difference. Of course, we can juggle the boundary condition to achieve more sensible answers but we still have unstoppable exponential growth predicted by the model.

R. WHITTINGTON

Hunters Way,
Bunkers Lane,
Rookley,
Isle of Wight, UK

1. Carpenter, F.L., Paton, D. & Hixon, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7259 (1983).
2. Butlin, R.K. *Nature* **312**, 313 (1984).

The record of global pollution in polar snow and ice

Eric W. Wolff & David A. Peel

British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK

Many attempts have been made to trace increases in global pollution from the record of impurities preserved in polar snow and ice layers. Although inadequate sampling and analysis techniques have hampered some of the studies, others have improved uniquely our understanding of the history of changes in airborne pollution.

POLAR ice sheets of Greenland and Antarctica have preserved a unique and dateable record of the past composition of the Earth's atmosphere. This record, which in places may extend for 500,000 years, contains information about climate and about environmental events such as volcanic activity. The more recent layers of ice preserve evidence for changes in large-scale air pollution¹⁻³.

Useful information is contained in three distinct chemical forms: (1) the stable isotope composition of the snow itself is determined by climatic factors; (2) airborne particulates, which include species such as heavy metals and sulphate, become trapped in falling snow or impacted on the snow surface; (3) as each snow layer sinks into the ice sheet, it compacts to form solid ice with enclosed air bubbles. These bubbles contain a sample of atmospheric gases.

Both particulates and gases will include pollutant species. Over large areas of the polar ice sheets where there is no melting, the chemical composition is preserved as the snow layers build up. Provided that whatever is incorporated in the ice sheet is representative of the atmosphere at a particular time, it is possible, by drilling into the ice, to obtain a record of past atmospheric composition.

In comparison with other sedimentary records, such as peat bogs and lake sediments, the ice sheets are particularly well resolved over the time-scale involved in pollution studies. A metre or more of snow accumulates each year at some sites. The polar regions are remote from local pollution sources, so that levels measured there represent large-scale averaged values. A comparison of present-day concentrations with natural, pre-cultural levels found in older snow may suggest both the extent and dispersal mechanisms of a pollutant.

Several gaseous and particulate pollutant species have been studied, mainly in Greenland and Antarctica. We describe here the information obtained and discuss the difficulties which have made much of the data unreliable.

Gases

Climatic change and gas content. Studies on polar ice cores have enabled the reconstruction of time series of climatic change in parallel with records of associated changes in the composition of the atmosphere. Comparison of the two records may help to elucidate whether they are linked causally. For example, evidence for the warming effect due to increasing levels of the 'greenhouse' gases such as CO₂ and CH₄ may be documented uniquely in the ice sheets. These effects can be viewed against the perspective of purely natural changes in the atmosphere which have occurred in earlier millennia.

Evidence for past temperature is recorded by fluctuations in the stable isotope content of the ice. The ratios of both ¹⁸O/¹⁶O and D/H are controlled primarily by the temperature at the time when the ice was deposited as snow. Seasonal changes can often be identified for several hundred years, allowing an ice

core to be dated stratigraphically and longer-term variations are indicative of past climate. Several deep ice cores from Antarctica⁴ and Greenland⁵ have shown a marked transition in the oxygen isotope content between 15,000 and 11,000 years ago. This was associated with an inferred temperature rise of ~7 °C at the end of the Wisconsin glaciation. A record extending to earlier ice ages should soon be available from central Antarctica where a borehole at Vostok station already penetrates through the Eem interglacial.

The past composition of gases in the atmosphere can be measured directly in air bubbles trapped in the ice. The age of this sample of the atmosphere, however, is not generally the same as that of the ice in which it is trapped. In the absence of melting, the air bubbles in ice become sealed off only at depth (typically 60–70 m). Therefore, the trapped air may be considerably younger than the ice in which it is enclosed. Because the pores do not close simultaneously, the bubbles at a particular depth contain air trapped at different times, which could restrict severely the resolution of a pollution record. At the best sites, with high accumulation and minimal melting, the age of 80% of the trapped air will fall within ~20 years⁶. Methane has been measured⁷ at one such site (Dye 3, Greenland), but a well-resolved record for CO₂ is still awaited.

Methane. Measurements in a core from Dye 3, Greenland⁷, showed a constant pre-industrial methane concentration of 0.70 parts per million by volume (p.p.m.v.), similar to other values reported in Greenland⁸ and Antarctica^{8,9}. Five samples from AD 1650 onwards showed⁷ an accelerating increase to 1.26 p.p.m.v. (sample date 1960), which agrees with some measured atmospheric concentrations in the 1960s (Fig. 1). There is, however, controversy¹⁰ as to how large atmospheric increases have been since 1948. Present-day concentrations are about 1.6 p.p.m.v. An earlier suggestion⁸ that low values from ancient ice could be the result of partial oxidation of CH₄ to CO within the ice is now discounted⁷, but it has been suggested that a diffusive loss mechanism could be responsible¹⁰.

If the increase seen in the ice cores is a genuine reflection of change in the atmosphere, it is more probably due to sources under human control, such as an increasing population of cattle and of rice paddy fields¹¹, rather than to increases in industrial emission rates. The increase could have been responsible for a greenhouse warming of 0.23 °C to date⁷. This conclusion is tentative until the true significance of the changes in methane concentration with ice depth can be elucidated.

Carbon dioxide. Most studies to date have examined several-thousand-year-old ice and have revealed a large and rapid increase from about 200 p.p.m.v. at the end of the Wisconsin glaciation to ~260 p.p.m.v. in the Holocene¹²⁻¹⁴. A series of rapid changes of CO₂ concentration¹⁵ (60 p.p.m.v. in time scales of the order of one century) in the Wisconsin glaciation are well correlated with climatic changes deduced from stable isotope records. It is still not clear whether the CO₂ changes are a cause

or an effect of climatic change, but oceanographic models have been developed to explain such a rapid change in CO_2 (refs 16, 17); CO_2 forcing could be an important amplifying factor in these fast climate variations¹⁵. These studies provide a better perspective for assessing the significance of modern increases in CO_2 concentrations.

The initiation of an increase in CO_2 concentrations in the late nineteenth and early twentieth centuries has been observed in an Antarctic ice core¹⁸. The few data available for ice only a few centuries old^{12-14,19} suggest that the pre-industrial CO_2 concentration was 260–270 p.p.m.v. (Fig. 1), a value in the range indicated by other, less direct, methods, but significantly smaller than values that have been used generally²⁰ for modelling recent CO_2 increases. Present-day air contains about 30% more CO_2 .

Early studies of CO_2 concentrations in air bubbles in ice were hampered both by contamination and inadequate extraction techniques. Two groups have developed dry extraction techniques that seem to give good agreement¹⁹ in a comparative study. A new technique²¹, however, gives consistent absolute values as high as 283 p.p.m.v. for the pre-industrial concentration, reopening the need for further comparative studies and suggesting that absolute values are still in doubt.

Simple models lead to an overall CO_2 -induced temperature rise of about 0.6 °C to date if the lower pre-industrial concentration is assumed. The increase would be reduced by nearly a half if this concentration was as high as 290 p.p.m.v.²². Natural temperature fluctuations would have obscured a change of this magnitude in instrumental records.

Particulate material

Unlike gases, particles are not well mixed in the global atmosphere. Whereas the atmospheric residence time for gases like CO_2 is typically several years, particles are transported mainly in the tropospheric circulation where the residence time lies in the range of weeks to a few months. Moreover, movement in the troposphere is restricted by circulation blocking systems at the Equator and at the polar convergence zones. Differences in snow composition between Greenland and Antarctica must therefore be expected because of the proximity of Greenland to other continents, sources of natural particulate material. In addition, 90% of industrial activity is in the Northern Hemisphere²³ and the heavily industrialized areas of North America and Europe are relatively close to Greenland. Antarctica is surrounded by the southern ocean and a strong zonal air circulation pattern which is shielded from the mid-latitude circulation by the polar front. The Greenland atmosphere is therefore likely to receive a greater input of both natural and anthropogenic particles than that of the Antarctic, reflected in the snow concentrations of many species.

Profiles of impurities in snow are generally interpreted simply in terms of temporal changes in the atmospheric aerosol. The validity of this interpretation depends on the mode of incorporation of particles into the snow²⁴. Fortunately, the few studies where air and snow concentrations have been compared have suggested that, within limits, there is a broadly linear relationship for a range of heavy metals and other trace elements²⁵⁻²⁸, but in only three of these studies²⁶⁻²⁸ (and for a few elements) were air and snow sampled simultaneously. In all cases, the data refer to averages over days or weeks; the heavy metal data in both air and snow in some studies²⁶ are probably erroneously high or given as upper limits²⁸. Therefore, the relationship between air and snow concentrations of many elements over a range of time scales requires considerably more investigation.

Heavy metals

Heavy metals are often defined as any metals with atomic number greater than that of iron. The elements we are concerned with have also been referred to as 'anomalously enriched elements' (AEEs) and as 'toxic metals and metalloids'. These elements are emitted industrially in large quantities²⁹, are dispersed widely in the environment³⁰ and many are potentially

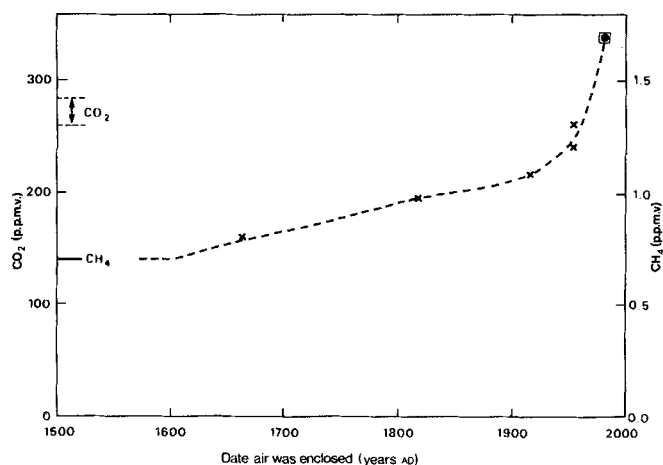


Fig. 1 Changes in atmospheric CH_4 (ref. 7) and CO_2 (refs 18–19, 21) with time, measured in air bubbles trapped in polar ice. Solid bar, pre-industrial Holocene average concentration of CH_4 ; two dashed bars, the range of pre-industrial values for CO_2 ; crosses, CH_4 concentrations measured in Greenland ice; square (CO_2) and circle (CH_4), present-day concentrations measured in the atmosphere.

harmful in very small concentrations. Although many data have been published, at the miniscule ($<1\text{--}200\text{ pg g}^{-1}$) concentrations in polar snow, successful analysis has often been hampered by contamination problems, hence, most of these data have been questioned^{31,32}.

It is important, therefore, to describe the conditions necessary for accurate and clean sampling and analysis and to pinpoint the reliable data. This should enable workers outside the immediate field to assess the validity of particular datasets.

Sampling and analysis. After a suitably remote sampling site has been found, all possible precautions must be taken against contamination, both in the field and laboratory. Sampling and clean-room techniques, suitable materials for tools and containers and scrupulous cleaning techniques have been described in detail elsewhere³¹⁻³⁸.

Samples have been taken from the snow surface^{28,34,39,40}, from trenches³³ and pits⁴¹⁻⁴⁵ (which at low snow accumulation sites have penetrated a century of snow^{41,42}), from an inclined shaft³³ and from blue ice areas, where ancient ice has been re-exposed^{31,33,46}. In addition, various types of drill have been used to recover ice cores. Most samples, and particularly cores, are inevitably dirty on the outside, so that a subsequent clean-up or sub-sampling is essential³². This may be more or less effective depending on the extent of penetration of contamination from the surface.

It is now considered essential that successive sub-samples from the outside surface to the centre of a sample should be analysed^{31,32,34} to show how far external contamination has penetrated to the centre. Only if a concentration plateau is found can data obtained from central areas of a sample be accepted. Using this approach, Ng and Patterson³⁵ showed that cores drilled using both thermal drills and electromechanical drills with drilling fluid are unsuitable for low-level heavy metal analysis, however stringent the subsequent clean-up. Satisfactory shallow cores have been obtained using a pre-cleaned polycarbonate auger³¹. The temperature at which cores are stored may influence crucially the rate of penetration of contamination. It is probable that cores drilled at sufficiently cold-field sites with an Al drill may also be suitable³⁴.

Pre-cleaning procedures, involving rinsing off the outer layers of a sample with pure water⁴³, have been used on solid ice but are probably unsuitable^{32,47}. Two dry methods have been used satisfactorily: (1) successive veneers are chiselled from the sample of firn or ice until an uncontaminated central area remains^{31,35,46}; (2) a PTFE tool is used to remove a small sub-sample from the centre of a firn sample³⁴.

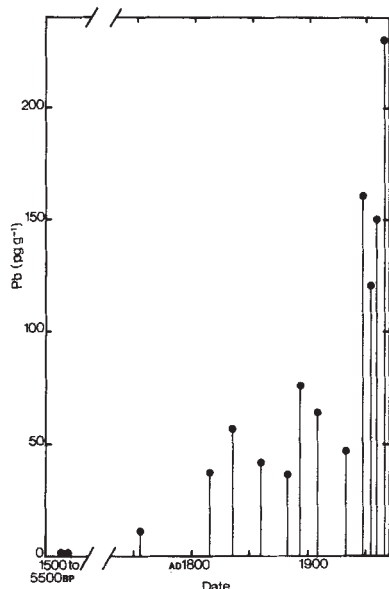


Fig. 2 Lead concentrations in Greenland snow³³.

Several analytical approaches, including isotope dilution mass spectrometry (IDMS), have been used for Pb analysis^{31,33,35}. This method is highly sensitive and absolute, but is also time-consuming, expensive and requires many stages of wet-chemistry. Anodic stripping voltammetry^{39,48,49} (ASV) and flameless atomic absorption spectrometry (FAAS) following a pre-concentration have been used for a range of metals. Pre-concentration onto tungsten wires^{26,34,50} uses smaller samples (~30 g) and is faster than evaporative pre-concentration^{40-42,44-46,51,52}, but cannot be used below pH 3. Instrumental neutron activation analysis (INAA) has been used mainly for Hg analysis⁵³⁻⁵⁵.

Analyses have been carried out under a range of conditions, from full acid digestion^{41,42} to analyses at pH 3 (refs 26, 34, 39, 48). The probable form of industrial metals reaching the polar regions suggests that most metals will be analysed even under the milder conditions³⁴, but care must be taken when comparing analyses from different methods. This is particularly important for older snow, where a high proportion of the metal may be in natural aluminosilicate based material.

Despite all precautions, procedural blanks must be determined³². Only a few workers^{31,34,35,46} have measured exactly how much contamination is added at each stage of their procedure and reported profiles showing how far superficial contamination has penetrated towards the centre of a sample. Without such a record, other data are hard to assess.

Lead in Greenland and the Arctic. In a pioneering study, Murozumi *et al.*³³ analysed for Pb in various samples from Greenland (Fig. 2). A few of the samples may contain contamination, but despite some criticism of the interpretation^{43,56-57} the results have been largely unchallenged.

The low value for old ice³³ has since been confirmed at Camp Century where a core sample several thousand years old contained 1.4 pg g⁻¹ at the centre³⁵. This value probably still includes some contamination.

Average concentrations in modern snow³³ in the range 150-250 pg g⁻¹ have been confirmed at sites in Greenland^{28,40,43}, Alaska⁵⁸ and the eastern Arctic Ocean³⁹. Lead (in common with several other elements) exhibits a distinct winter maximum, corresponding to a period when atmospheric conditions favour the input of pollutants from Europe³⁹. Much higher reported concentrations⁵⁹ must be erroneous⁶⁰. The suggestion⁶¹ that it is the lower values^{33,35} which are wrong runs strongly against the consensus. Data for old ice obtained by thermal drilling⁴³ must now be dismissed³⁵.

In summary, Pb in Northern Hemisphere polar snow seems

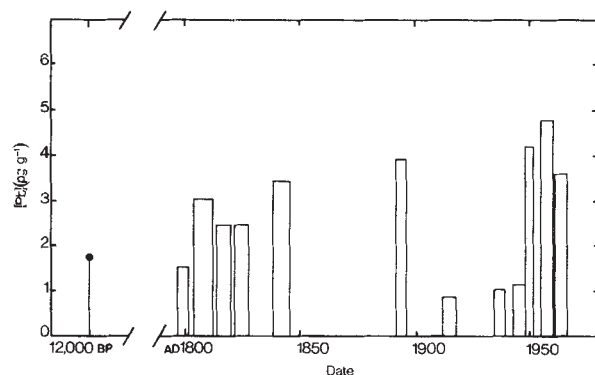


Fig. 3 Lead concentrations in snow from Greater Antarctica^{31,35}. Three depths where limited melting was observed have been excluded.

to have an average present-day concentration of ~200 pg g⁻¹, with a late summer minimum of about 13 pg g⁻¹ (ref. 39). By contrast, several-thousand-year-old snow contained at most 1.4 pg g⁻¹ Pb. The exact time-course of this increase (see Fig. 2) has yet to be confirmed. There is some evidence³³ that corresponding increases in the air may have been somewhat smaller than changes observed in the ice, because of preferential deposition of particles containing anthropogenic lead.

Lead in Antarctica. Successive veneers from cores drilled with a polycarbonate drill and from a block of blue ice from Greater Antarctica have been analysed recently³¹. Some samples showed signs of limited melting and the remaining centre-core data are shown in Fig. 3. For some of the samples, cross-core contamination profiles clearly showed a plateau; for others the situation was less clear. Assuming the data in Fig. 3 are all correct, however, it is not clear whether the variations are due to as yet unidentified natural sources such as volcanism³¹, or whether they represent a small anthropogenic increase. The value for ancient ice^{31,35} may still include contamination, as the across-core profile does not show any plateau when plotted on a logarithmic scale³⁴. In a study supported also by full procedural blanks and contamination profiles, a value of 6 pg g⁻¹ was reported for modern snow³⁴. Although this was at a much higher accumulation site in the Antarctic Peninsula, this is in good agreement with the data of Bortron and Patterson³¹. This recent work invalidates much of the previous Pb data^{26,33,41,42,44,45,48,52,59,62-65} for Antarctica.

Although the newer data are a great improvement on earlier studies, uncertainty in some values, particularly for ancient ice, means that any anthropogenic increase could still lie anywhere in the range 1-40-fold³⁴ with smaller values probably still favoured³². The best available data are summarized in Table 1. The much larger modern increases in Pb concentrations in Greenland snow compared with Antarctic snow reflect the fact that 90% of anthropogenic emissions occur in the northern hemisphere²³.

Other heavy metals in Greenland and the Arctic. (1) Hg. The most recent data⁵⁵ gave values of 2-15 pg g⁻¹, with no trend between 1927 and 1971. These values are considerably lower than earlier ones^{54,66,67}, but may still include contamination³². (2) Cd, Cu, Ni. For these elements, the best available data are from a study of near surface snow in the eastern Arctic Ocean³⁹, where annually averaged values of Cd (5 pg g⁻¹), Cu (97 pg g⁻¹) and Ni (195 pg g⁻¹) were found and values observed which were one order of magnitude lower than this in snow that fell in late summer. These lower values were suggested to be the maximum possible pre-industrial contribution³⁹. Some similar concentrations of Cd and Cu have been reported in Greenland^{40,43,44}, but higher reported values probably result from contamination^{66,68}. There are no reliable data for these elements in old snow and ice. (3) Zn. For surface snow, Zn concentrations around 200 pg g⁻¹ have been reported^{40,43-44}. Herron *et al.*⁴³ reported 77 pg g⁻¹ in pre-industrial snow in Greenland, but this certainly

includes some contamination. (4) Ag. Modern Ag concentrations in the range 6–13 pg g⁻¹ have been reported^{40,44} from central Greenland, though values as low as 1.2 pg g⁻¹ were found in mid-summer at Dye 3 (ref. 28). There are no reliable values for ancient ice.

Other heavy metals in Antarctica. (1) Hg. At Mizuho station, increases in Hg concentrations from <1 pg g⁻¹ in some older samples (~900 years old) to more than 20 pg g⁻¹ in some recent samples were reported^{63,64}. The elevated concentrations, however, seem to extend up to two centuries into the past. Taken together with the fact that associated Pb concentrations are too high, we believe these Hg data are erroneous. (2) Cd, Cu, Zn. In a study with full procedural blanks and cross-core contamination profiles, values for surface snow of Cd (0.26 pg g⁻¹), Cu (1.8 pg g⁻¹) and Zn (3.3 pg g⁻¹) were obtained³⁴. Although these are for a few weeks' snow at a high accumulation Antarctic Peninsula site, these values are probably reasonably representative³⁴. All earlier reported concentrations^{26,41–42,44–45,59,62,65} are now thought to be wrong^{32,34}. The most recently published data for ancient ice⁴⁶ give values an order of magnitude larger than the new modern snow values. However, contamination has probably influenced these data³⁴. We conclude that there are no reliable data for these elements in ancient Antarctic ice. (3) Ag. Concentrations in the range 0.3–8 pg g⁻¹, with no obvious increase with time since 1894 (ref. 69), have been reported for a number of Antarctic sites^{69,70}. Slightly higher values have also been reported^{41,42,45}.

Enrichment factors. Measurements of heavy metal concentrations in air and in both fresh and marine water environments in remote areas of the earth^{71–73} have shown often large factors of enrichment compared with elements of exclusively crustal origin. Studies on Antarctic air⁷⁴ have generally paralleled the findings from snow, showing a tendency for increasing degrees of enrichment with increasing elemental volatility.

Examination of all the data now available for polar snow, however, suggests that contamination has played an important role in elevating the crustal enrichment factors (EFs) of most, if not all, of the heavy metals determined in snow (Table 1). The most reliable studies are now suggesting that some elements (for example, Cu, Pb, Zn) in ancient ice have EFs between 1 and 10.

Very much reduced EFs have also been reported recently for heavy metals in Antarctic air²⁷. We suggest this may be the trend in other remote environments and that the concept of large natural enrichments may no longer be required to explain newer less-contaminated data.

Other particulates

Radioactive debris. High concentrations of β -radioactivity can be identified in snow layers of the polar ice caps which corresponded to known nuclear weapons tests in the atmosphere⁷⁵. Natural tritium is estimated to have had an activity (at the time of precipitation) of 26 ± 3 TU at the South Pole in the period

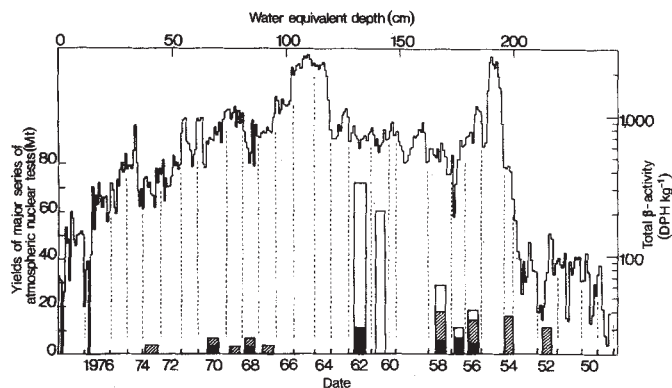


Fig. 4 Record of total β -radioactivity at South Pole 1950–1978 (ref. 25). The bars show the yields of major series of atmospheric nuclear tests⁸¹. Latitudes of emission: solid 2° N–30° S (UK 1956–58, USA 1962, France 1968 onwards); hatched 10° N–40° N (USA 1952–58, China 1967 onwards); white 75° N and other USSR sites.

1939–42 (ref. 76). At its peak in 1966, the activity was 2,000 TU (ref. 77). Detailed chronologies have been described from both hemispheres, for total β -radioactivity^{25,78–80} (Fig. 4) (due mainly to ⁹⁰Sr and ¹³⁷Cs), and for tritium^{76,77,81–86} derived from bomb tests. Particular layers such as those corresponding to the Castle tests in 1954 have been used extensively as reference horizons for the dating of snow^{87,88}. The profiles have also been used to determine other meteorological parameters, such as the stratospheric residence time^{25,80} of nuclear debris.

Specific isotopes have been studied to investigate individual events. The burn-up, in 1964, of a satellite (SNAP-9A) is clearly seen in ²³⁸Pu profiles in both hemispheres^{85,89–90}. Similarly, increased ratios of ²³⁴U/²³⁸U at Dye 3, Greenland, were attributed⁹¹ to debris from Cosmos 954, which burnt up over Canada in 1978. Tests of nuclear weapons carried out at oceanic sites in the 1950s appear in profiles of ³⁶Cl (produced from seawater chlorine) at Dye 3 (ref. 92). In 1960, ³⁶Cl concentrations were 100-fold greater than pre-bomb levels.

Several other artificial radionuclides occur in both hemispheres^{85–86}, interpreted in terms of particular types and sizes of weapon tested. A comparison of profiles from the northern and southern hemispheres helps to pinpoint the location of the tests.

Organochlorine compounds. We assume that DDT, polychlorinated biphenyls (PCBs) and related compounds were non-existent in pre-cultural air and snow.

The first values reported for DDT in Antarctic snow⁹³ were much higher than those reported later^{94,95} (0.1 – 4.0×10^{-12} g g⁻¹ for DDT and metabolites). The ratio of DDT concentrations in snow to those in penguins was consistent with those expected for animals in equilibrium with their environment⁹⁴. Concentra-

Table 1 Most reliable concentrations of heavy metals reported in polar snows

Location	Concentrations (pg g ⁻¹)						
	Cadmium	Copper	Lead	Mercury	Nickel	Silver	Zinc
Modern Greenland or Arctic Ocean snow	5 (ref. 39)	97 (ref. 39)	200 (ref. 33)	10* (ref. 55)	195 (ref. 39)	10* (ref. 40)	200 (ref. 43)
Ancient Greenland ice	?	?	1.4* (ref. 35)		?	?	<77* (ref. 43)
Modern Antarctic snow	0.26 (ref. 34)	1.8 (ref. 34)	5 (refs 31, 34)	10* (ref. 63)	?	0.5 (ref. 69)	3.3 (ref. 34)
Ancient Antarctic ice	?	?	1.2* (ref. 31)	2* (ref. 63)	?		?
Concentration expected with EF = 1† and Al = 1 ng g ⁻¹	0.0024	0.67	0.15	0.001	0.91	0.0009	0.85

* Upper limit, value very uncertain. † Average crustal compositions from ref. 71. 1 ng g⁻¹ Al is reasonable for many Antarctic sites; Enrichment factor

$$(EF) = \frac{[X]_{\text{snow}} / [Al]_{\text{snow}}}{[X]_{\text{crust}} / [Al]_{\text{crust}}}$$

where [] are the elemental concentrations in snow and in average crustal rock. ? Values unknown. For ancient Greenland ice, values of at most 0.4 pg g⁻¹ (Cd), 13 pg g⁻¹ (Cu) and 22 pg g⁻¹ (Ni) can be inferred from modern seasonal minima³⁹.

Table 2 Sulphate concentrations in ancient and modern snows

Site	Analytical technique	Ref.	Range of years sampled	Ancient snow	Range of years sampled	Modern snow
				Concentration of SO ₄ ²⁻ (ng g ⁻¹)		Concentration of SO ₄ ²⁻ (ng g ⁻¹)
GREENLAND						
Camp Century	A	117	pre-1930	81	1964-65	200
	A	118	Holocene average	81	—	—
	C	104	1788-1800	42	1976-77	120
Milcent	A	43	1170-1886	81	1972-73	206
North central	C	104	1442-50	80	1973-77	130
Dye 3	A	119	1252-1820	56	1970-71	205
	A	66	1807-1915	74	1966-71	126
(Dye 3 area)	B	120	1680-1845	47	1974-75	97
	C	104	1539-1891	22	1975-81	86
ANTARCTIC						
Dome C	B	111	1880-1920	85	1970-78	80
Dumont d'Urville to Dome C	B	121			1971-75	Average 71
Vostok	A	122			1971	120
Byrd Station	A	117	1724	78	—	—
	A	118	Holocene average	61	—	—
	C	104	—	—	Modern	45
Ross Ice Shelf C-16	C	104	1590-1598	84*	1971-77	61*
Ross Ice Shelf Q-13	C	104	1590-1598	98*	1973-77	103*
South Pole	B	121			1968-74	73
	C	104			Modern	45
	C	100			1959-69	72

* Includes marine aerosol sulphate. Analytical methods: A, barium chloranilate; B, lead electrode; C, ion chromatography.

tions of PCBs reported at the same time were probably derived partially from activities at a nearby base⁹⁵. Recent measurements⁹⁶ of DDT and metabolites ($9\text{--}17 \times 10^{-15} \text{ g g}^{-1}$), PCBs ($0.16\text{--}1.0 \times 10^{-12} \text{ g g}^{-1}$) and HCHs (hexachlorocyclohexane isomers) ($1.5\text{--}4.9 \times 10^{-12} \text{ g g}^{-1}$) in snow from near Syowa and Mizuho stations showed no significant difference between 1960 and 1980 snow. The new, lower DDT concentration leaves the true value in some doubt.

As DDT has been largely phased out in the Northern Hemisphere, but not yet in the Southern Hemisphere, a comparison of detailed profiles from the two hemispheres should reflect the circulation of this chemical.

Acidity. Studies of the acidity of polar snowfall are of value for assessing the relative importance of natural and anthropogenic emissions to the global atmospheric cycle. Evidence for the background acidity of precipitation has been reviewed in detail elsewhere⁹⁷. The mineral acids H_2SO_4 and HNO_3 are major impurities in polar snow and are probably dominant in the interior of Antarctica, remote from primary aerosol sources. H^+ (or pH), sulphate and nitrate have all been measured in polar snow.

Nevertheless, there are still few reliable data on the acidity of polar snow, mainly because the concentrations of soluble impurities are very low and measurements are sensitive to contamination by traces of alkaline dusts and gases⁹⁸ during sample handling and analysis. It is only since the advent of ion chromatography⁹⁸ and acid titration techniques⁹⁹ that satisfactory ionic balances have been achieved¹⁰⁰. Hence, caution must be taken when comparing absolute values obtained by different methods. pH measurements are intrinsically insensitive to monitoring absolute changes in acidity and difficult to measure accurately in samples of very low ionic strength.

A very rapid method for profiling changes in acidity along ice cores records the electrical conductivity between two electrodes scratched along the surface of the ice¹⁰¹⁻¹⁰³. The acidity is generally related simply to the observed conductivity and the method has been used successfully to identify specific acid horizons linked with explosive volcanic events. Eruptions in the Northern Hemisphere are particularly well recorded in the Greenland Ice Sheet¹⁰², where it has been possible to reconstruct the chronology of major volcanic events during the past 10,000 years. These studies have also given much new information on the type and intensity of the eruptions. This method does not, however, provide an accurate absolute measurement of the acid concentration and is therefore not suitable for studying trends in acidity such as those linked with industrial emissions.

Table 2 extends Herron's listing¹⁰⁴ of sulphate data for both ancient and modern polar snows. It excludes data that are unrepresentative because they cover a few days snowfall only^{28,105}, appear contaminated¹⁶² or were obtained from sites

close to the sea¹⁰⁶⁻¹⁰⁸. Taking the data as a whole it seems that there has been an approximately threefold increase in non-marine sulphate concentration in Greenland snow this century. For the same period, Herron¹⁰⁴ has reported a doubling in concentration of nitrate at three Greenland sites. These results are qualitatively consistent with an observed decrease of pH during the last 26 years of 0.007 pH units per year in snow from the Canadian Arctic¹⁰⁹. The trend probably parallels increasing long-range transport of products of combustion of fossil fuels. Note that no systematic increase in acidity since 1850 was observed in meltwater conductivity measurements¹¹⁰ or from scratch conductivity measurements¹⁰² on an ice core from Crete, Greenland. This discrepancy may reflect some residual analytical difficulties and underlines the need for a continuous time series of absolute acidity data from an uncontaminated, remote site.

The sparser time-series data for sulphate in Antarctica show no systematic increasing trend although there are large fluctuations during the last 100 years which seem to be linked with major volcanic eruptions in the Southern Hemisphere¹¹¹. Background sulphate concentrations in central Antarctica almost doubled in the years following the eruptions of Krakatoa (1883) and Mt Agung (1963) and volcanic activity probably accounts for one third of the total sulphate content of Antarctic snow during the last century. These findings have been confirmed by acid titration⁹⁹ of strong acids in snow at both Dome C and the South Pole where no trend was observed over the last 20 years^{100,112}. The background acidity is dominated by H_2SO_4 which is probably mainly a product of the oxidation of sulphur compounds generated biologically over the surrounding oceans^{97,111}. There is, however, no satisfactory explanation for the significant concentrations of HNO_3 . A connection with solar activity and supernova has been proposed^{113,114} but the data supporting of this hypothesis have been questioned^{97,104,115}. The negligible influence of long-range transport of industrially derived acids to Antarctica probably reflects the fact that 94% of anthropogenic sulphur is emitted in the Northern Hemisphere¹¹⁶.

Conclusions

Studies on polar snow and ice during the last 15 years have revealed their enormous potential for improving our understanding of the history of large-scale changes in airborne pollution. Evidence for a twofold increase in the concentration of methane in the atmosphere has been documented over the period since 1650, but its interpretation in terms of increases in sources under human control is still controversial. A pre-industrial CO_2 concentration of 260 p.p.m.v. has been determined, compared with a present-day concentration of 340 p.p.m.v. The

pre-industrial value is significantly smaller than that assumed until recently in models of CO₂-induced warming of the atmosphere.

Concentrations of sulphate and nitrate in Greenland snow have increased threefold and twofold respectively in the last century. In contrast, no recent increases have been detected for these acidic species in Antarctica. This reflects the geographical pattern of industrial emissions and the restricted movement of tropospheric aerosols between the hemispheres. Similarly, much larger surface snow concentrations of heavy metals are found in Greenland than in Antarctica. Apart from data for Pb in Greenland, however, which show an increase of at least 100-fold since 800 BC, there are no reliable data for heavy metals in

ancient ice in either hemisphere, so anthropogenic trends cannot yet be estimated.

Despite sustained effort, only recently have experimental methods improved sufficiently to monitor reliably the extremely small concentrations of pollution transported to these remote areas. It is now possible to identify the broad characteristics of change between natural, pre-industrial conditions and those of the present day. Future work of increased accuracy will be required in order to develop detailed time-series from the polar ice sheets. These are required urgently to test models of the global circulation of pollution.

We thank our colleagues at the British Antarctic Survey and Dr C. Boutron, for many constructive criticisms.

1. Hammer, C. U. in *Atmospheric Chemistry* (ed. Goldberg, E. D.) 119–134 (Springer, 1982).
2. Herron, M. M. & Herron, S. L. *Hydrol. Sci. J.* **28**, 139–153 (1983).
3. Peel, D. A. *New Scientist*, **98**, 476–483 (1983).
4. Lorius, C., Merlivat, L., Jouzel, J. & Pourchet, M. *Nature* **280**, 644–648 (1979).
5. Dansgaard, W. et al. *Science* **218**, 1273–1277 (1982).
6. Schwander, J. & Stauffer, B. *Nature* **311**, 45–47 (1984).
7. Craig, H. & Chou, C. C. *Geophys. Res. Lett.* **9**, 1221–1224 (1982).
8. Robbins, R. C., Cavanagh, L. A., Salas, L. J. & Robinson, E. J. *Geophys. Res.* **78**, 5341–5344 (1973).
9. Khalil, M. A. K. & Rasmussen, R. A. *Chemosphere* **11**, 877–883 (1982).
10. Ehhalt, D. H., Zander, R. J. & Lamontagne, R. A. *J. geophys. Res.* **88**, 8442–8446 (1983).
11. Rasmussen, R. A. & Khalil, M. A. K. *J. geophys. Res.* **86**, 9826–9832 (1981).
12. Delmas, R. J., Ascencio, J.-M. & Legrand, M. *Nature* **284**, 155–157 (1980).
13. Oeschger, H., Stauffer, B., Neftel, A., Schwander, J. & Zumbund, R. *Ann. Glaciol.* **3**, 227–232 (1982).
14. Neftel, A., Oeschger, H., Schwander, J., Stauffer, B. & Zumbund, R. *Nature* **295**, 220–223 (1982).
15. Stauffer, B., Hofer, H., Oeschger, H., Schwander, J. & Siegenthaler, U. *Ann. Glaciol.* **5**, 160–164 (1984).
16. Sarmiento, J. L. & Toggweiler, J. R. *Nature* **308**, 621–624 (1984).
17. Siegenthaler, U. & Wenk, Th. *Nature* **308**, 624–626 (1984).
18. Raynaud, D. & Barnola, J. M. *Ann. Glaciol.* **5**, 224 (1984).
19. Barnola, J. M., Raynaud, D., Neftel, A. & Oeschger, H. *Nature* **303**, 410–413 (1983).
20. Lorius, C. & Raynaud, D. in *Carbon Dioxide, a Text on Current Views and Developments in Energy/Climate Research* (eds Bach, W. et al.) 145–176 (Dordrecht, Reidel, 1983).
21. Moor, E. & Stauffer, B. *J. Glaciol.* **30**, 358–361 (1984).
22. Siegenthaler, U. & Oeschger, H. *Ann. Glaciol.* **5**, 153–159 (1984).
23. Robinson, E. & Robbins, R. C. *Emission, concentrations and fate of particulate atmospheric pollutants* (Am. Petr. Inst. Publ. 4070, Washington DC, 1971).
24. Junge, C. E. in *Isotopes and Impurities in Snow and Ice* No. 118, 63–77 (IAHS, Grenoble, 1977).
25. Pourchet, M., Pinglot, F. & Lorius, C. *J. geophys. Res.* **88**, 6013–6020 (1983).
26. Peel, D. A. & Wolff, E. W. *Ann. Glaciol.* **3**, 255–259 (1982).
27. Dick, A. L. & Peel, D. A. *Ann. Glaciol.* **7** (in the press).
28. Davidson, C. I., Chu, L., Grimm, T. C., Nasta, M. A. & Qamoos, M. P. *Atmos. Envir.* **15**, 1429–1437 (1981).
29. Nriagu, J. O. *Nature* **279**, 409–411 (1979).
30. *Ninth Rep. Royal Commission Enq. Pollution* paras 8.13–8.14 (HMSO, London, 1983).
31. Boutron, C. F. & Patterson, C. C. *Geochim. cosmochim. Acta* **47**, 1355–1368 (1983).
32. Boutron, C. F. in *Toxic Metals in the Air* (eds Davidson, C. I. & Nriagu, J. O.) (Wiley, New York, in the press).
33. Murozumi, M., Chow, T. J. & Patterson, C. C. *Geochim. cosmochim. Acta* **33**, 1247–1294 (1969).
34. Wolff, E. W. & Peel, D. A. *Ann. Glaciol.* **7** (in the press).
35. Ng, A. & Patterson, C. C. *Geochim. cosmochim. Acta* **45**, 2109–2121 (1981).
36. Boutron, C. *Analyt. chim. Acta* **106**, 127–130 (1979).
37. Patterson, C. C. & Settle, D. M. *National Bureau of Standards Spec. Publ.* **422**, 321–351 (1976).
38. Patterson, C. C. *J. Ass. Official analyt. Chem.* (1982).
39. Mart, L. *Tellus* **35B**, 131–141 (1983).
40. Boutron, C. *Geochim. cosmochim. Acta* **43**, 1253–1258 (1979).
41. Boutron, C. & Lorius, C. *Nature* **277**, 551–554 (1979).
42. Boutron, C. *J. geophys. Res.* **85**, 7426–7432 (1980).
43. Herron, M. M., Langway, C. C. Jr., Weiss, H. V. & Cragin, J. H. *Geochim. cosmochim. Acta* **41**, 915–920 (1977).
44. Boutron, C. Z. *für Gletscherkunde und Glazialgeologie*, **17**, 79–113 (1981).
45. Boutron, C. *Atmos. Envir.* **16**, 2451–2459 (1982).
46. Boutron, C., Leclerc, M. & Risler, N. *Atmos. Envir.* **18**, 1947–1954 (1984).
47. Boutron, C. F. & Batifol, F. M. *Ann. Glaciol.* **7** (in the press).
48. Landy, M. P. & Peel, D. A. *Nature* **291**, 144–146 (1981).
49. Landy, M. P. *Analyt. chim. Acta* **121**, 39–49 (1980).
50. Wolff, E. W., Landy, M. P. & Peel, D. A. *Analyt. Chem.* **53**, 1566–1570 (1981).
51. Boutron, C. & Martin, S. *Analyt. Chem.* **51**, 140–145 (1979).
52. Boutron, C. & Martin, S. *J. geophys. Res.* **85**, 5631–5638 (1980).
53. Herron, M. M. et al. in *Isotopes and Impurities in Snow and Ice* No. 118, 98–102 (IAHS, Grenoble, 1977).
54. Weiss, H. V., Koide, M. & Goldberg, E. D. *Science* **174**, 692–694 (1971).
55. Appelquist, H., Jensen, K. O., Sevel, T. & Hammer, C. *Nature* **273**, 657–659 (1978).
56. Dickson, E. M. *Science* **177**, 536–538 (1972).
57. Robinson, J. *Ecotoxicol. environ. Safety* **5**, 24–37 (1981).
58. Hinkley, T. K. in *USGS Open File Report* No. 78-701, 179–180 (1978).
59. Jaworowski, Z., Bysiek, M. & Kownacka, L. *Geochim. cosmochim. Acta* **45**, 2185–2199 (1981).
60. Patterson, C. C. *Geochim. cosmochim. Acta* **47**, 1163–1168 (1983).
61. Jaworowski, Z., Bysiek, M. & Kownacka, L. *Geochim. cosmochim. Acta* **47**, 1169–1175 (1983).
62. Gjessing, Y. T. & Gjessing, E. T. *Vatten* **3**, 233–237 (1973).
63. Murozumi, M., Nakamura, S. & Yoshida, Y. *Mem. Nat. Inst. Polar Res. Spec. Issue* **7**, 255–263 (1978).
64. Yoshida, Y. & Murozumi, M. *Ant. Record (Japan)* **59**, 26–29 (1977).
65. Miklishanskiy, A. Z., Yakovlev, Y. V., Savelyev, B. V. & Zuyev, A. P. *Geochem. Int.* **1980**, 152–158 (1981).
66. Weiss, H., Bertine, K., Koide, M. & Goldberg, E. D. *Geochim. cosmochim. Acta* **39**, 1–10 (1975).
67. Carr, R. A. & Wilkiss, P. E. *Science* **181**, 843 (1973).
68. Weiss, H. V., Koide, M., Bertine, K. K. & Goldberg, E. D. in *Isotopes and Impurities in Snow and Ice* No. 118, 103–107 (IAHS, Grenoble, 1977).
69. Warburton, J. A., Linkletter, G. O., Owens, M. S., Young, L. G. & Anderson, A. V. *J. rech. Atmos.* **9**, 119–126 (1975).
70. Warburton, J. A., Molenaar, J. V., Owens, M. S. & Anderson, A. *Pure appl. Geophys.* **118**, 1130–1144 (1980).
71. Taylor, S. R. *Geochim. cosmochim. Acta* **28**, 1273–1285 (1964).
72. Nriagu, J. O. (ed.) *The Biogeochemistry of Lead in the Environment* (North Holland, Amsterdam, 1978).
73. Nriagu, J. O. (ed.) *Cadmium in the Environment* (Wiley-Interscience, New York, 1980).
74. Maenhaut, W., Zoller, W. H., Duce, R. A. & Hoffman, G. L. *J. geophys. Res.* **84**, 2421–2431 (1979).
75. Picciotto, E. & Wilgain, S. *J. geophys. Res.* **68**, 5965–5972 (1963).
76. Jouzel, J., Merlivat, L., Mazaudier, D., Pourchet, M. & Lorius, C. *Geophys. res. Lett.* **9**, 1191–1194 (1982).
77. Jouzel, J., Merlivat, L., Pourchet, M. & Lorius, C. *Earth planet. Sci. Lett.* **45**, 188–200 (1979).
78. Croaz, G., Langway, C. C. Jr & Picciotto, E. *Res. Rept.* **208** (US Army Cold Regions Research and Engineering Laboratory, 1966).
79. Lambert, G., Ardouin, B., Sanak, J., Lorius, C. & Pourchet, M. in *Isotopes and Impurities in Snow and Ice* No. 118, 146–158 (IAHS, Grenoble, 1977).
80. Pourchet, M. & Pinglot, J. F. *Geophys. Res. Lett.* **6**, 365–367 (1979).
81. Carter, M. W. & Moghissi, A. A. *Health Physics* **33**, 55–71 (1977).
82. Merlivat, L., Ravoire, J., Vergnaud, J. P. & Lorius, C. *Earth planet. Sci. Lett.* **19**, 235–240 (1973).
83. Merlivat, L., Jouzel, J., Robert, J. & Lorius, C. in *Isotopes and Impurities in Snow and Ice* No. 118, 138–145 (IAHS, Grenoble, 1977).
84. Hebert, D. & Frohlich, K. *ZfM-Mitteilungen* **51**, 87–100 (1982).
85. Koide, M., Michel, R., Goldberg, E. D., Herron, M. M. & Langway, C. C. Jr *Earth planet. Sci. Lett.* **44**, 205–223 (1979).
86. Koide, M., Michel, R., Goldberg, E. D., Herron, M. M. & Langway, C. C. Jr. *Nature* **296**, 544–547 (1982).
87. Aristarain, A. J. & Delmas, R. J. *J. Glaciol.* **27**, 371–379 (1981).
88. Peel, D. A. & Clausen, H. B. *J. Glaciol.* **28**, 43–55 (1982).
89. Cutter, G. A., Bruland, K. W. & Risebrough, R. W. *Nature* **279**, 628–629 (1979).
90. Koide, M., Goldberg, E. D., Herron, M. M. & Langway, C. C. Jr *Nature* **269**, 137–139 (1977).
91. Koide, M. & Goldberg, E. D. *Earth planet. Sci. Lett.* **65**, 245–248 (1983).
92. Elmore, D. et al. *Nature* **300**, 735–737 (1982).
93. Peterle, T. J. *Nature* **224**, 620 (1969).
94. Peel, D. A. *Nature* **254**, 324–325 (1975).
95. Risebrough, R. W., Walker, W. II, Schmidt, T. T., de Lappe, B. W. & Connors, C. W. *Nature* **264**, 738–739 (1976).
96. Tanabe, S., Hidaka, H. & Tatsukawa, R. *Chemosphere* **12**, 277–288 (1983).
97. Delmas, R. J. & Gravenhorst, G. in *Proc. EEC workshop on Acid Deposition* (eds Beilke, F. & Elshout, A. J.) 82–107 (Reidel, Berlin, 1984).
98. Legrand, M., De Angelis, M. & Delmas, R. J. *Analyt. chim. Acta* **156**, 181–192 (1984).
99. Legrand, M. R., Aristarain, A. J. & Delmas, R. J. *Analyt. Chem.* **54**, 1336–1339 (1982).
100. Legrand, M. R. & Delmas, R. J. *Atmos. Envir.* **18**, 1867–1874 (1984).
101. Hammer, C. U. *J. Glaciol.* **25**, 359–372 (1980).
102. Hammer, C. U., Clausen, H. B. & Dansgaard, W. *Nature* **288**, 230–235 (1980).
103. Hammer, C. U. *J. phys. Chem.* **87**, 4099–4103 (1983).
104. Herron, M. M. *J. geophys. Res.* **87**, 3052–3060 (1982).
105. Delmas, R. J., Barnola, J. M. & Legrand, M. *Ann. Glaciol.* **3**, 71–76 (1982).
106. Lorius, C., Baudin, G., Cittanova, J. & Platzler, R. *Tellus* **21**, 136–148 (1969).
107. Vileskii, V. D. & Koroleva, N. I. *Antarktika* **12**, 94–101 (1973).
108. Aristarain, A. J., Delmas, R. J. & Briat, M. *J. geophys. Res.* **87**, 11004–11012 (1982).
109. Koerner, R. M. & Fisher, D. *Nature* **295**, 137–140 (1982).
110. Hammer, C. U. *Nature* **270**, 482–486 (1977).
111. Delmas, R. & Boutron, C. *J. geophys. Res.* **85**, 5645–5649 (1980).
112. Delmas, R. J., Aristarain, A. & Legrand, M. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 104–105 (SNSF, Oslo, 1980).
113. Rood, R. T., Sarazin, C. L., Zeller, E. J. & Parker, B. C. *Nature* **282**, 701–703 (1979).
114. Parker, B. C., Zeller, E. J. & Gow, A. J. *Ann. Glaciol.* **3**, 243–248 (1982).
115. Risbo, T., Clausen, H. B. & Rasmussen, K. L. *Nature* **294**, 637–639 (1981).
116. Cullis, C. F. & Hirschler, M. M. *Atmos. Envir.* **14**, 1263–1278 (1980).
117. Koide, M. & Goldberg, E. D. *J. geophys. Res.* **76**, 6589–6596 (1971).
118. Cragin, J. H., Herron, M. M., Langway, C. C. Jr & Klouda, G. in *Polar Oceans* (ed. Dunbar, M. J.) 617–631 (Arctic Institute of North America, Calgary, 1977).
119. Cragin, J. H., Herron, M. M. & Langway, C. C. Jr *Res. Rept.* **341**, US Army Cold Regions Research and Engineering Laboratory (1975).
120. Busenberg, E. & Langway, C. C. Jr *J. geophys. Res.* **84**, 1705–1709 (1979).
121. Delmas, R. & Boutron, C. *Atmos. Envir.* **12**, 723–728 (1978).
122. Doronin, A. N. *Inf. Byull. sov. Antarkt. Eksped.* **91**, 62–69 (1975).

Lower mantle heterogeneity, dynamic topography and the geoid

Bradford H. Hager, Robert W. Clayton, Mark A. Richards, Robert P. Comer* & Adam M. Dziewonski*

Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA

Density contrasts in the lower mantle, inferred using seismic tomography, drive viscous flow; this results in kilometres of dynamically maintained topography at the core-mantle boundary and at the Earth's surface. The total gravity field due to interior density contrasts and dynamic boundary topography predicts the longest-wavelength components of the geoid remarkably well. Neglecting dynamic surface deformation leads to geoid anomalies of opposite sign to those observed.

THE LONGEST-WAVELENGTH components of the Earth's gravity field have been known for more than two decades from satellite geodesy^{1,2} but there have been few observational constraints on the density anomalies causing them. Developments in seismology and theoretical geodynamics, until recently separate branches of geophysics, give new insight into this classical problem.

Seismological studies of lateral heterogeneity have determined the longest-wavelength components of seismic velocity variations in the lower mantle³⁻⁶. These velocity variations presumably would be proportional to density variations if both were the result of temperature differences associated with mantle convection. But these inferred density variations are, in fact, negatively correlated with the observed geoid, apparently contradicting the convection hypothesis.

Geodynamic theories of the Earth's gravity field show, however, that a negative correlation is not unexpected in a convecting Earth, because of the counteracting effects of dynamically supported surface deformations on the gravity field⁷. For example, hot upwellings in the mantle cause uplift of the Earth's surface and of the core-mantle boundary. This dynamically maintained topography has a strong effect on the gravity field. For a given interior density structure, the total gravity field (including the effects of surface deformation) depends on the viscosity structure and on the presence or absence of chemical stratification in the mantle^{8,9}.

Our main conclusion is that the longest-wavelength components of the residual geoid can be predicted from the seismically-inferred density contrasts in the lower mantle using dynamic response functions for mantle convection with a moderate increase in viscosity with depth. The inferred long-wavelength temperature differences are of the order of $\pm 20^\circ\text{C}$, with hot lower mantle underlying hotspot provinces at the Earth's surface. The inferred deformation of the core-mantle boundary is ~ 3 km.

It is possible to predict accurately the longest-wavelength geoid using the lower mantle seismic results and mantle flow models which include the effects of possible chemical stratification at the 670-km seismic discontinuity or with models which assume mantle-wide flow. Mantle-wide flow and stratified flow can, in principle, be distinguished on the basis of predicted dynamic topography at the Earth's surface. This dynamic topography, of the order of 1 km, is of opposite sign for the two styles of flow. Although incomplete knowledge of global dynamic topography and possible masking effects of dynamic topography because of upper mantle heterogeneities complicate interpretation, approximate agreement between model predic-

tions and observations of topographic anomalies is suggestive of mantle-wide convection.

Seismic velocity structure

We describe here two independent methods to invert similar International Seismological Centre (ISC) P-wave travel-time residual data sets for the velocity structure of the lower mantle. The first method, Dz (from Dziewonski⁴), involves expressing velocity perturbations δV in terms of smooth functions—Legendre functions in radius r and spherical harmonics in colatitude θ and longitude φ :

$$\delta V(r, \theta, \varphi) = \sum_{k=0}^K \sum_{l=0}^L \sum_{m=0}^l f_k(r) \cdot ({}_kA_l^m \cos m\varphi + {}_kB_l^m \sin m\varphi) P_l^m(\cos \theta)$$

Some 500,000 travel-time residuals for teleseismic arrivals from 5,000 earthquakes were used in an iterative least-squares procedure to derive the coefficients A and B for $K=4$ and $L=6$ for 245 coefficients.

The second topographic method (CC, from Clayton and Comer^{5,6}) is an iterative back projection of travel-time residuals along ray paths. The mantle is discretized into 29 spherical shells, each 100-km thick, with each shell divided into 1,676 cells ~ 500 km on a side. At every iteration, the slowness ($1/V$) perturbation, δS_i , in the i th cell is incremented by

$$\frac{\sum_{k=1}^M (l_{ki} \delta t_k / L_k)}{\mu + \sum_{k=1}^M l_{ki}}$$

where, for the k th ray, δt_k is the residual relative to the preceding iteration, L_k is the total ray length, l_{ki} is the ray length in the cell (which may be zero), M is the total number of rays and μ is a damping parameter. The procedure converges to the smallest δS_i in the sense of a Euclidean norm, where the cells intersected by more rays are given greater significance than others, that minimizes

$$\sum_{k=1}^K \left[\left(\delta t_k - \sum_{i=1}^N l_{ki} \delta S_i \right)^2 / L_k \right]$$

Here, N is the total number of cells, K is the total number of rays and δt_k is a residual relative to the (spherically symmetric) starting model.

Each inversion method has advantages and disadvantages. Method Dz gives a direct global inversion for the long wavelength features of interest for comparison with the geoid and provides a resolution matrix with formal uncertainties. It is a low-order polynomial fit to sparse data, allowing model

* Present addresses: Department of Earth and Space Sciences, State University of New York, Stony Brook, New York 11794, USA (R.P.C.); Department of Geological Sciences, Harvard University, Cambridge, Massachusetts 02138, USA (A.M.D.).

excursions in regions not well constrained by data and not allowing rapid radial variation of the model. Method CC allows determination of finer structure and provides a simple visualization of those regions not covered by rays, making it easy to see regions that are not constrained by observations. It does not, however, provide formal uncertainties. Because it is expanded ultimately in terms of spherical harmonics for comparison with the gravity field, it shares the problem of how best to interpolate into regions not well constrained by observations. We derive the spherical harmonic expansion coefficients used here with a least-squares fit to the blocks in a given layer.

At this stage, the primary limitations for both approaches are the non-uniform distribution of ray coverage brought about by both the uneven distribution of sources and receivers and noise in the travel-time data. It is encouraging that the two methods agree in the longest-wavelength components of velocity anomalies ($l=2-3$), although the agreement is poor at shorter wavelengths ($l \geq 4$). We confine our interpretation to those wavelengths where the methods agree, justified in part by the fact that at these wavelengths the seismic models can predict the geoid if dynamic effects are taken into account properly.

Generation of geoid anomalies

Although on short time scales the mantle behaves as an elastic solid, it is recognized commonly that on geological time scales it behaves as a viscous fluid, responding to stresses by slow creeping flow. Less commonly recognized is that the flow resulting from interior density contrasts (such as those inferred from studies of variations in seismic velocity) leads to deformation of the Earth's surface, the core-mantle boundary and any other boundary in chemical composition which may exist in the mantle⁷⁻⁹. This boundary deformation occurs on the same time scale as postglacial rebound, short compared with the time it takes for the position of the interior density pattern to change appreciably⁸. Thus, boundary deformation can be taken to occur instantaneously from the standpoint of mantle convection.

The total geoid anomaly observed at the surface is the sum of the opposing effects of the interior density contrasts driving the flow and the mass anomalies caused by boundary deformations resulting from the flow. The amplitudes of the boundary deformations, and hence the sign and magnitude of the total geoid anomaly, depend on the distribution of viscosity with depth and the presence or absence of chemical stratification⁸⁻¹⁰.

We use spherical harmonics to describe the gravity field, as well as the seismic velocity heterogeneities. Given the spherical harmonic coefficient of density perturbation as a function of radius, $\delta\rho_l^m(r)$, the total gravitational potential U_l^m is given by the integral:

$$U_l^m = \frac{4\pi\gamma a}{2l+1} \int_c^a G_l^m(r) \delta\rho_l^m(r) dr \quad (3)$$

Here, γ is the gravitational constant, a is the radius of the Earth, c the radius of the core and $G_l^m(r)$ is the dynamic response function or kernel⁸⁻¹⁰. This kernel includes the contributions of the boundary deformations caused by the flow induced by $\delta\rho_l^m(r)$, as well as from the driving density contrast itself.

We calculate dynamic response functions using models that treat the mantle as an incompressible newtonian fluid with a spherically-symmetric viscosity structure. The m dependence of G is then degenerate and G is a function of l and r only. Both the viscosity structure and the intrinsic density profile due to differences in chemical composition can be layered. Phase changes in a layer of uniform composition are assumed to have negligible effect on the flow field, as phase changes by themselves do not present a barrier to radial flow. (Any horizontal density contrasts arising from phase changes are, at least in principle, part of the seismically-inferred density field.)

Free-slip boundary conditions are applied at the Earth's surface and at the core-mantle boundary. At boundaries between layers of the same chemical composition but differing viscosities, velocities and tractions are continuous. At boundaries between

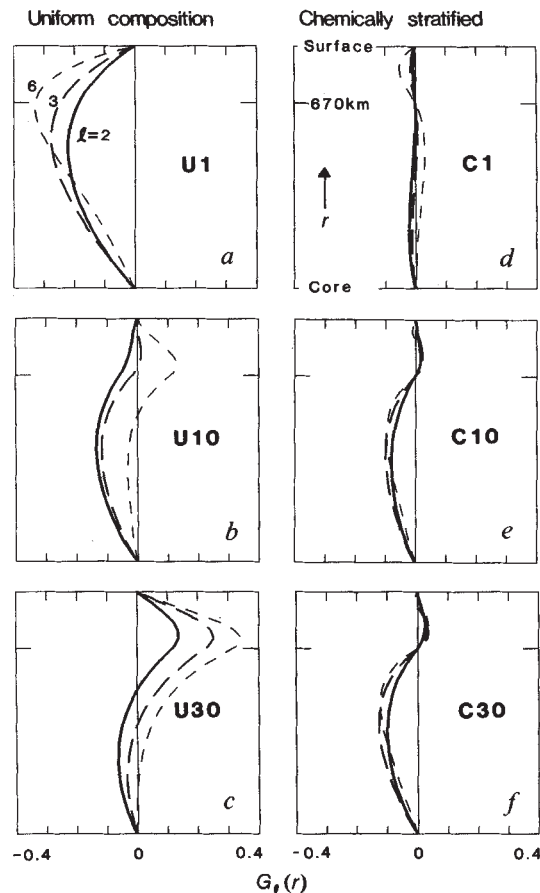


Fig. 1 Dynamic response functions for surface density contrasts of spherical harmonic degrees 2 (solid lines), 3 (long dashes), and 6 (short dashes) plotted against radius for six Earth models. Models U, left, have uniform composition which permits mantle-wide flow; models C, right, have a chemical discontinuity at 670-km depth, causing stratification into separate upper and lower mantle flow systems. Models in the top row have uniform viscosity; those in the middle row have a factor of 10 viscosity increase below 670 km. This viscosity increase is a factor of 30 for models in the bottom row.

layers of differing chemical compositions, radial velocities are set to zero, prohibiting radial flow between layers. Other velocity components and tractions are continuous, leading to separate flow systems with shear coupling across the boundary.

Our assumption of a newtonian, spherically-symmetric viscosity distribution is done for the sake of mathematical tractability and to provide a simple model of the relevant physics. The viscosity of mantle rocks depends on both temperature and stress¹¹; lateral variations in viscosity are expected to accompany the observed lateral heterogeneity in seismic velocity. The existence of tectonic plates (which translate as nearly rigid blocks) with deformation concentrated in zones of weakness at plate boundaries, demonstrates that lateral viscosity variations are important in the Earth. Our primary justification for using such simple models is that they are adequate to explain the longest-wavelength components of the geoid. The main effect of long-wavelength lateral variations in viscosity is to excite flow at shorter wavelengths, leaving the geoid signature at the long wavelengths relatively unaffected¹².

Examples of dynamic response functions $G_l(r)$ for six simple viscous Earth models are shown in Fig. 1 for spherical harmonic degrees 2, 3 and 6. U1, U10 and U30 in the left column represent mantle models of uniform composition with a ratio of lower-mantle to upper-mantle viscosity of 1, 10 and 30 respectively. (Although flow velocities depend on the actual value of viscosity chosen, stresses and geoid anomalies depend only on the relative distribution of viscosity.) Models C1, C10 and C30 in the right

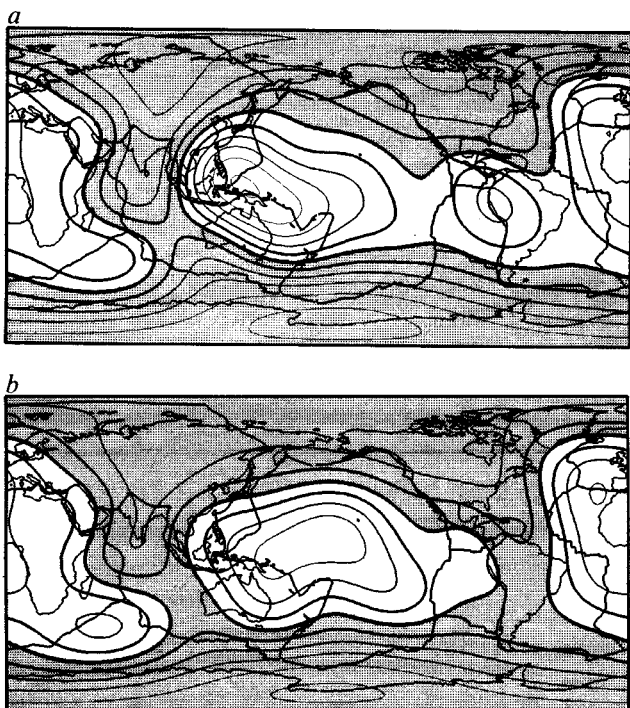


Fig. 2 *a*, The observed geoid¹³ for $l=2-6$ referred to the hydrostatic figure of the Earth. Geoid lows are shaded and the contour interval is 20 m. In all our maps, we show plate boundaries and continents for reference. *b*, The residual geoid for $l=2-6$ obtained by subtracting the effects of subducted slabs¹⁰. Lows are shaded; the contour interval is 20 m.

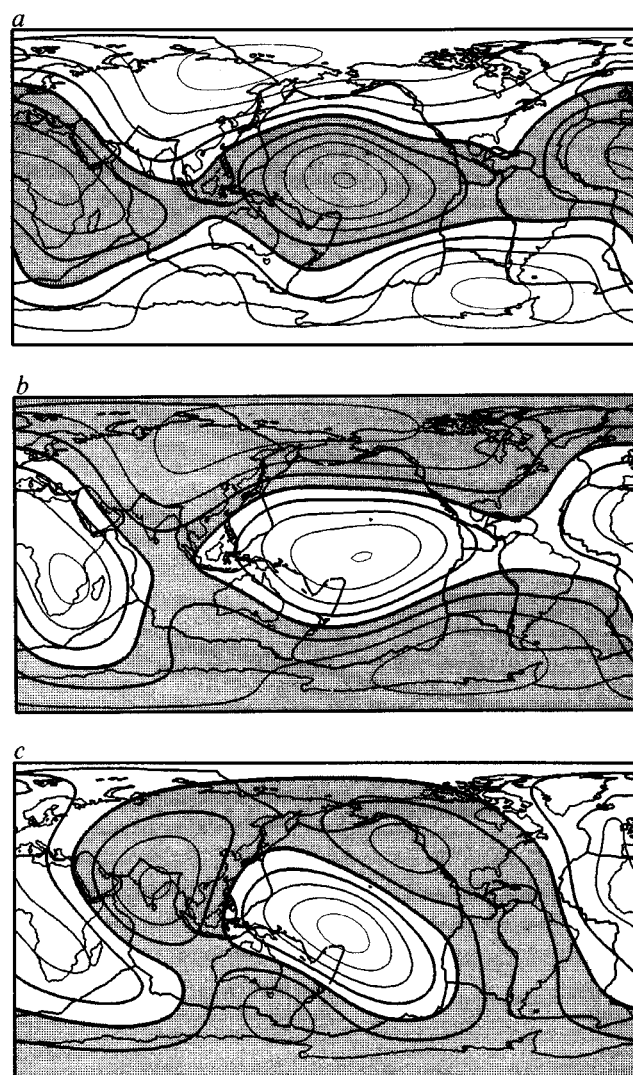


Fig. 3 Geoids calculated by applying dynamic response functions, such as those in Fig. 1, to the seismic models. *a*, CC-S, a static model (boundary deformation is not included); contour interval, 50 m. *b*, CC-U10, for a uniform composition model with a factor of 10 increase in lower mantle viscosity. *c*, Dz-U10, as *b*; contour intervals, 20 m. Lows are shaded.

column are for models with an intrinsic chemical density contrast at 670-km depth, with separate flow systems above and below the 670-km seismic discontinuity. The viscosity structure for these models is identical to that in the corresponding row for the uniform composition models.

For a given density contrast, the magnitude and sign of the resulting geoid anomaly in a dynamic Earth depends on the viscosity structure. For example, in model U1, the opposing gravitational effect of deformation of the upper boundary overwhelms that of the interior density contrast itself, leading to a negative geoid anomaly for a positive density contrast. For model U30, because of the strong increase in viscosity with depth, most of the deformation occurs at the core-mantle boundary, further away from the observer than the density anomaly driving the flow. The effect of the density contrast dominates and the response function is more positive. The net gravitational field of a dynamic Earth is a small number determined by the difference of larger, nearly counter-balancing quantities. The sign of the field depends on which of the effects is dominant. The anomaly may depend also on the depth of the convecting system, with chemically stratified systems leading usually to smaller geoid anomalies for a given density anomaly. Observations of both the geoid and seismic velocity heterogeneities over a range of wavelengths place constraints on the variation of mantle viscosity and the depth of mantle convection. Observation of dynamic surface deformation provides a complementary constraint.

Results

Figure 2*a* shows a recent long-wavelength ($l=2-6$) geoid¹³, referred to the hydrostatic equilibrium figure¹⁴, superimposed on a map including plate boundaries. There is a strong association of geoid highs with subducted slabs^{10,15-17} but there is much variation, particularly at the longest wavelengths, not associated with the present plate configuration.

This can be seen more clearly in Fig. 2*b*, a residual geoid determined by subtracting from the observed geoid slab effects obtained using a mantle flow model in conjunction with a model of the density structure of subducted slabs¹⁰. This flow model assumes mantle-wide flow and an upper mantle with a viscosity 0.01 times that of the lower mantle. Other more empirical studies give similar residual geoids^{16,17}. The residual geoid has a pool-ball like pattern dominated by antipodal highs over north-west Africa and the central Pacific. The highs are in regions shielded from subduction for substantial parts of the recent geological past and might be expected to represent hotter than average mantle^{18,19}. Indeed, there is a strong correlation between long-wavelength residual geoid highs and hotspots^{16,17}.

These geoids can be compared with model geoids calculated using the lower-mantle seismic models. Velocity and density perturbations are assumed to be directly proportional and the seismically inferred density anomalies are convolved with the response functions shown in Fig. 1, as well as those for a static (rigid) Earth. If dynamic surface deformation is ignored (Fig. 3*a*, model CC-S, for static), the geoid anomaly matches the pattern, in general, but is of opposite sign to the observed or residual geoids. Thus, we must consider the gravitational effects of surface deformation.

For uniform composition and viscosity (model U1, not shown) there is general agreement between the calculated and the residual geoids, but the high-frequency component has too great an amplitude in the calculated geoid and the best fitting proportionality constants for degrees 2 and 3 differ substantially. For model U10, with uniform composition and a lower mantle viscosity increased by a factor of 10, there is an excellent correlation between the calculated and residual geoids at longest wavelengths for either model CC (Fig. 3b) or model Dz (Fig. 3c). This agreement breaks down for model U30, which has a factor of 30 increase in lower mantle viscosity.

For the chemically layered model C1, the agreement is poor. Models C10 and C30 predict geoids similar to each other (seen by comparing the response functions in Fig. 1e and f). Either provides a good match to the residual geoid, almost as good as model U10. The general agreement between models U10 and C10 (and C30) is expected from the similarity of their response functions at $l=2-3$, Fig. 1b and 1e.

Correlation coefficients for degrees 2-6 are given in Table 1 for the observed geoid, the residual geoid and models Dz-U10 and CC-U10. For both models, the agreement at $l=2-3$ is excellent, with the best fitting proportionality constant about $(4 \text{ km s}^{-1})/(\text{g cm}^{-3})$, close to that predicted for likely mantle compositions and observed for variations in velocity and density with radius near the top of the lower mantle²⁰. The positive correlation is consistent with the convection hypothesis, with both high velocities and high densities associated with cold material.

The dynamically maintained topography at the Earth's surface and at the core-mantle boundary for degrees 2-3, calculated for model CC-U10, is shown in Fig. 4. Both the core-mantle boundary and the surface are upwarped in regions of low seismic velocity and downwarped in regions of high velocity (and presumably high density). These dynamically induced deformations of the top and bottom of the convecting system represent substantial mass anomalies. Their effects on the geoid are so large that they overwhelm the effects of the internal density contrasts driving the flow, thereby determining the sign of the geoid.

The total geoid is correlated positively with dynamic surface deformations and negatively with seismically inferred density anomalies. This confirms the need to consider the effects of dynamic surface deformation in calculating geoid anomalies; neglecting it leads to an error in sign as well as in magnitude.

Because the gravitational response functions for flow models U10 and C10 are so similar, it is not possible to discriminate between these two models using only the longest-wavelength geoid predicted from our present models of seismic heterogeneity in the lower mantle. However, these two models predict quite different values for the dynamically maintained topography at the Earth's surface. The dynamic surface topography predicted for the chemically stratified model CC-C10 is quite similar in pattern to that for model CC-U10, Fig. 4a. However, it is of opposite sign due to the shear coupling of flow between the upper and lower mantle in a chemically stratified mantle, and is a factor of three lower in amplitude. If it were possible to strip out the isostatic effects of crustal thickness variations and the dynamic effects of upper mantle heterogeneity to isolate the dynamic component of surface topography due to lower-mantle density anomalies, it would be possible to distinguish between these models.

An accurate determination of the long-wavelength component of dynamic topography requires a knowledge of the global distribution of crustal densities and thickness variations, beyond the scope of this report. If we assume, however, that depth or elevation anomalies in similar tectonic provinces are due to dynamic surface deformation, we can begin to compare Fig. 4a with some observations. Comparing ancient shield regions which have been unaffected by crustal thickening in the last billion years (and might, therefore, have been eroded to the same elevation), we find that Africa stands higher than the Canadian or Siberian shields by several hundred metres⁸. There is a broad

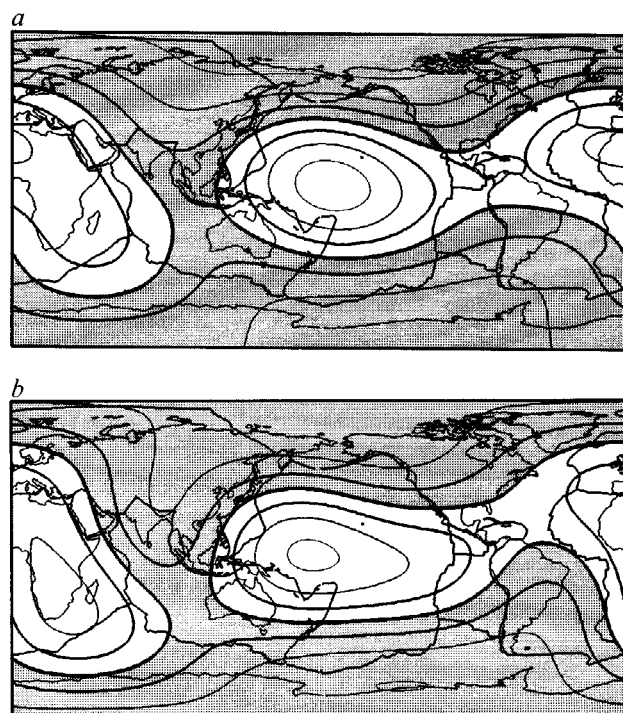


Fig. 4 Calculated topography for $l=2-3$, *a*, at the surface and *b*, core-mantle boundary for the model CC-U10, which matches the geoid. Surface topography was calculated assuming a density contrast of 2.3 g cm^{-3} (appropriate for sub-oceanic regions). Topography at the core-mantle boundary is calculated for a density contrast of 4.5 g cm^{-3} across the interface. Lows are shaded. Contour intervals: *a*, 250 m; *b*, 500 m.

region of anomalously shallow (250-750 m) sea floor in the Western Pacific and an anomalously deep ($> 500 \text{ m}$) region south of Australia²¹. Africa is surrounded mainly by anomalously shallow seafloor. Comparing long-wavelength model predictions with a spotty and possibly aliased distribution of points on a map is dangerous. But we find that the agreement of our uniform-composition model predictions with what is known of the distribution of residual topographic anomalies, and the disagreement with the predictions of the chemically stratified model, suggest mantle-wide convection.

Temporal variations in large amplitude, dynamically-maintained surface topography would contribute to epeirogeny and eustatic sea-level changes. For example, changes in the positions of continents and ocean basins relative to the underlying convection pattern itself could explain easily the range of eustatic sea-level changes observed. Records of 'eustatic' sea level from different continents would depend on their positions relative to the long-wavelength mantle flow pattern.

Dynamically maintained topography at the core-mantle boundary has a total excursion of $\sim 3 \text{ km}$ for either model. Such large topography has been suggested as influencing the geodynamo²² and may be important in coupling nutation of the mantle to the core. The dynamic topography for degrees 2-3 is a substantial fraction of the 9-km topography due to the hydrostatic ellipticity of the core-mantle boundary. Higher-degree terms are expected to increase the dynamic topography substantially.

Discussion

There is remarkable agreement between the residual geoid and the geoid calculated using the dynamic response functions for models U10 or C10 using either seismic model. The agreement between the predictions of either seismic model and the geoid is comparable with the agreement between the two seismic models. Using the seismic models (and the dynamic slab

model¹⁰) we can explain more than 80% of the variance in the observed $l = 2-3$ geoid. This agreement is unlikely to be due to chance.

There are certainly large heterogeneities in the upper mantle which might be expected *a priori* to mask the signal from the lower mantle. For example, several seismic studies have reported a degree 2 velocity anomaly in the upper mantle which correlates well with the degree 2 geoid²³⁻²⁵. It is remarkable that the fit of the geoid is so good, as we have neglected upper mantle contributions. The longest-wavelength dynamic response functions for our successful models (U10 or C10), however, are quite small in the upper mantle. The $l = 2-3$ geoid is not as sensitive to density anomalies there as it is to those in the lower mantle; the effects of heterogeneities in the transition zone are compensated almost totally by surface deformation.

The seismic results show that at degrees 2-3, geoid lows are associated with fast lower mantle whereas highs are associated with slow lower mantle. We suggest that these velocity variations are caused by temperature differences of convective origin. This interpretation is supported by a very high correlation ($r > 0.99$) between the degree 2 distribution of hotspots and the radially-averaged degree 2 lower mantle seismic velocity anomalies. This supports the hypotheses that geoid lows result from areas where subduction has cooled the mantle¹⁹ and that shielding of the mantle by continents results in geoid highs and hotspots¹⁸. Mantle-wide flow provides a simple explanation for the fact that hotspot provinces at the surface tend to overly the long-wavelength expression of hot lower mantle. The relatively higher viscosity we find for the lower mantle may help to explain the long time scales inferred to be associated with these phenomena.

We have confined our analysis here to fairly simple viscosity distributions that, nonetheless, have allowed us to explain most of the variance in the longest-wavelength components of the geoid. The viscosity models presented are not unique; viscosity values should be regarded as weighted averages over the two

layers used. We can provide even better matches to the geoid using more complicated viscosity models, including, for example, a low viscosity channel beneath a high viscosity layer representing a lithospheric lid. The best two-layer viscosity model used in fitting the seismic results to the residual geoid has a smaller viscosity contrast between the upper and lower mantle than the two-layer model used to isolate and remove slab effects¹⁰, consistent with the lithosphere or upper mantle having a lower effective viscosity at plate boundaries than in plate interiors. It is possible, however, using a uniform composition model with four viscosity layers, to achieve a self-consistent spherically symmetric model for both the inferred lower mantle density anomalies and subducted slabs (B.H.H., in preparation).

Finally, although the agreement at long wavelengths is impressive, the agreement at $l \geq 4$ between the predictions of these models and the geoid is not good. This is in part the result of oversimplification in the flow models, which contain only two layers and which assume a spherically symmetric effective viscosity. On the other hand, the disagreement between the seismic methods at $l \geq 4$ leads to the hope that improved seismic data (such as will be provided by the proposed Global Seismic Array) and new techniques for its analysis, as well as including results for upper mantle heterogeneity, will yield an understanding of the geoid at shorter wavelengths. The insights into mantle dynamics provided by even these relatively coarse seismic studies are extremely valuable. Further cooperative studies by seismologists and geodynamacists should lead to even better constraints on the dynamics of convection in the Earth's mantle.

We thank D. L. Anderson, C. Chase, R. J. O'Connell and D. J. Stevenson for helpful comments. This work was supported by NASA grants NAG5-315, NSG-7610, NAS5-27226, NSF grant EAR-8317623, by the Alfred P. Sloan Foundation (B.H.H.), and by a Bantrell Postdoctoral Fellowship (R.P.C.). Contribution no. 4065, Division of Geological and Planetary Sciences, California Institute of Technology.

Received 2 July; accepted 29 October 1984.

1. Kaula, W. M. *J. geophys. Res.* **68**, 473-484 (1963).
2. Guier, W. H. *Nature* **200**, 124-125 (1963).
3. Dziewonski, A. M., Hager, B. H. & O'Connell, R. J. *J. geophys. Res.* **82**, 239-255 (1977).
4. Dziewonski, A. M. *J. geophys. Res.* **89**, 5929-5952 (1984).
5. Comer, R. P. & Clayton, R. W. *J. geophys. Res.* (submitted).
6. Clayton, R. W. & Comer, R. P. *J. geophys. Res.* (submitted).
7. Pekeris, C. L. *Mon. Not. R. astr. Soc. Geophys. Suppl.* **3**, 343-367 (1935).
8. Richards, M. A. & Hager, B. H. *J. geophys. Res.* **89**, 5987-6002 (1984).
9. Ricard, Y., Fleitout, L. & Froidevaux, C. *Annls Geophys.* **2**, 267-286 (1984).
10. Hager, B. H. *J. geophys. Res.* **89**, 6003-6015 (1984).
11. Ashby, M. F. & Verrall, R. A. *Phil. Trans. R. Soc. A* **288**, 59-95 (1977).
12. Richards, M. A. & Hager, B. H. *EOS* **65**, 1092 (1984).

13. Lerch, F. J., Klosko, S. M. & Patch, G. B. *LAGEOS (GEM-L2)* (NASA Technical Memorandum, 84986, 1983).
14. Nakiboglu, S. M. *Phys. Earth planet. Inter.*, **28**, 302-311 (1982).
15. Kaula, W. M. *The Nature of the Solid Earth* (ed. Robertson, E. G.) 385-405 (McGraw-Hill, New York, 1972).
16. Chase, C. G. *Nature* **282**, 464-468 (1979).
17. Crough, S. T. & Jurdy, D. M. *Earth planet. Sci. Lett.* **48**, 15-22 (1980).
18. Anderson, D. L. *Nature* **297**, 391 (1982).
19. Chase, C. G. & Sprowl, D. R. *Earth planet. Sci. Lett.* **62**, 314-320 (1983).
20. Bass, J. D. & Anderson, D. L. *Geophys. Res. Lett.* **11**, 237-240 (1984).
21. Crough, S. T. *A Rev. Earth planet. Sci.* **11**, 165 (1983).
22. Malin, S. R. C. & Hide, R. *Phil. Trans. R. Soc. A* **306**, 281 (1982).
23. Masters, G., Jordan, T. H., Silver, P. G. & Gilbert, F. *Nature* **298**, 609-613 (1982).
24. Woodhouse, J. & Dziewonski, A. M. *J. geophys. Res.* **89**, 5953-5986 (1984).
25. Nakanishi, I. & Anderson, D. L. *Nature* **307**, 117-121 (1984).

Segmental distribution of bithorax complex proteins during *Drosophila* development

Philip A. Beachy*, Stephen L. Helfand† & David S. Hogness*

Departments of *Biochemistry and †Biological Sciences, Stanford University School of Medicine, Stanford, California 94305, USA

The Ubx and bxd transcription units comprise a single functional domain in the bithorax complex of Drosophila melanogaster. The segmental distributions and nuclear localization of proteins encoded by the Ubx unit have been determined by immunofluorescence staining with antibodies raised against a fusion protein containing Ubx coding sequences. Wild-type and mutant distributions are consistent with a model in which the protein-coding functions of the domain derive from the Ubx unit and are regulated by the bxd unit.

THE metazoan genome faces a unique set of demands in its task of directing the complex process of multicellular development. Not only must the characteristics of individual cells be specified, but also the size, shape and other higher-order characteristics of the multicellular structures they constitute. In *Drosophila melanogaster*, the body plan of both larva and fly is characterized

by a series of segments whose identities vary according to position along the anterior-to-posterior axis of the animal. The role genes play in the specification of these identities has been defined, in part, by homoeotic mutations that produce specific intersegmental transformations of identity by causing embryonic founder cells at one position to adopt the fates reserved normally

for those at another position. The genes defined by these mutations are thus presumed to mediate the developmental decisions that differentiate one segment from another.

The identities of the second and third thoracic segments (T2, T3) and all eight of the abdominal segments (A1-8) depend on a cluster of homoeotic genes known as the bithorax complex (BX-C)^{1,2}. This complex is divided into three functional domains³. We concentrate here on the Ubx (Ultrabithorax) domain, which occupies the left or proximal one-third of the BX-C DNA (Fig. 1) and acts in the specification of the T2, T3 and A1 identities. More precisely, the multiple functions of the Ubx domain specify the identities of the four adjacent anterior (a) and posterior (p) compartments (T2p, T3a, T3p and Ala) that mark the anterior end of the segmental range of BX-C expression. The Ubx domain is defined by five classes of recessive loss-of-function mutations that can be divided into two groups^{1,3-10}. One group consists of four classes (*abx*, *bx*, *bxd* and *pbx*) which serve to define individual, compartment-specific functions of the domain because each causes the homoeotic transformation of a different subset of the four compartments. By contrast, the fifth or *Ubx* class of mutations *cis*-inactivates all functions in the set defined by the other classes and thereby defines that set as a single functional domain; thus, *Ubx* mutations do not complement any of the other mutations and cause homoeotic transformations corresponding to the sum of those induced by the other classes.

Given this *cis*-linkage of domain functions, it is curious that the molecular mapping summarized in Fig. 1 reveals a division of the domain into two non-overlapping, commonly oriented transcription units: the ~75-kilobase (kb) *Ubx* unit, which is virtually coextensive with the region delineated by the sites of *Ubx* mutations and which includes the *abx* and *bx* loci, and the ~25-kb *bxd* unit, which is similarly coextensive with the *bxd* sites and overlaps the *pbx* mutations. The problem of how *Ubx* mutations can *cis*-inactivate all functions of the domain, while the other classes of mutations *cis*-inactivate specific subsets of these functions, is a fundamental aspect of the more general problem of how the bithorax complex specifies segmental identities. As a step towards a better definition of these problems we have begun an investigation of the proteins encoded by the *Ubx* unit.

Open reading frames in the *Ubx* unit

First, we summarize the relevant aspects of an analysis of the *Ubx* RNAs to be described elsewhere (Fig. 2 and R. B. Saint, P.A.B., D. Peattie, P. Harte and D.S.H., in preparation). Three major size classes of *Ubx* RNAs appear in the first 8 h of embryogenesis: a transient 4.7-kb poly(A)⁻ RNA that disappears near the end of this time and two poly(A)⁺ RNAs of 3.2 and 4.3 kb found at all subsequent stages of development. The 3.2-kb class first appears at ~3 h, slightly after the 4.7-kb and before the 4.3-kb RNAs. Figure 2 shows that the 3.2-kb RNAs contain a long AUG-initiated open reading frame (ORF) generated by two exons from the 5' and 3' ends of the *Ubx* unit and two internal 51-base pair (bp) micro-exons. Sequence analysis of two cDNA clones shows that either of two donor sites located 27-bp apart may be used to splice the 5' exon to the first micro-exon. The resulting ORF Ia and ORF Ib contain, respectively, 380 and 389 codons in the same frame and have the capacity to encode two polypeptides of relative molecular mass (*M_r*) 39,000 and 40,000, which differ only by the insertion of nine amino acids between residues 247 and 248 of the shorter polypeptide.

Several indicators other than significant length suggest that these ORFs encode *Ubx* proteins: their codon usage agrees well with codon frequencies tabulated for known *D. melanogaster* structural genes (K. Burtis, unpublished); their 3'-exon regions contain a homoeo box (a ~180-bp conserved sequence found in several *D. melanogaster* homoeotic and segmentation genes, and in the genomes of many other species¹¹⁻¹⁴); the 5'-exon region common to both ORFs is highly conserved among *D. melanogaster*, *Drosophila funebris* and *Musca domestica* (D.

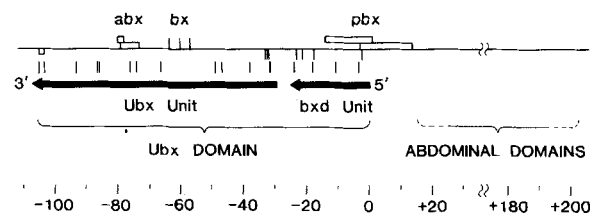


Fig. 1 Molecular map of the Ubx domain in the BX-C. The location of DNA lesions associated with recessive mutations in the *Ubx* domain are indicated in kb relative to the *Ubx* and *bxd* transcription units. The mutation sites are taken from Bender *et al.*²⁰ and from M. Akam and D.S.H. (unpublished data). Positions of the *Ubx* and *bxd* units are taken, respectively, from R. B. Saint, P.A.B., D. Peattie, P. Harte and D.S.H. (in preparation) and unpublished data of M. Goldschmidt-Clermont, D. Peattie, H. Lipshitz and D.S.H. The breakpoints of *Ubx* chromosomal rearrangements that split the BX-C are indicated by the vertical lines in the row just above the solid arrow denoting the orientation and extent of the *Ubx* unit. Four pseudo-point *Ubx* mutations (three small deletions and an insertion) are indicated by an open box (3' end) and three closely spaced vertical slashes (5' end) just below the long horizontal line representing the chromosomal DNA; they are also shown in Fig. 2 at higher resolution. The breakpoints of *bxd* rearrangements and the positions of *bxd* pseudo-point mutations are similarly indicated above the solid arrow delineating the *bxd* unit. The *abx* and *pbx* deletions are denoted by open boxes above the line whereas the sites of *bx* insertional mutations are indicated by vertical slashes.

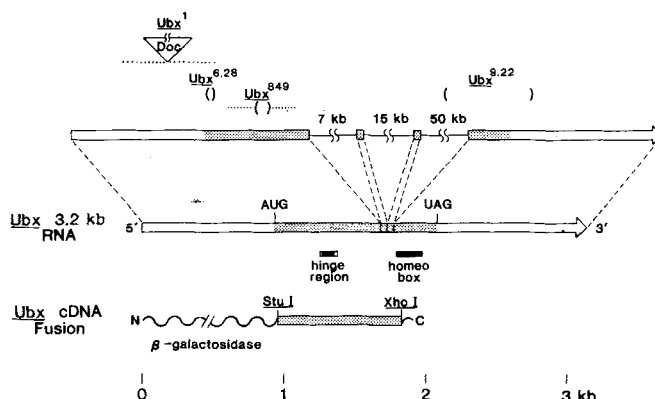


Fig. 2 The *Ubx* open reading frame used to construct the *Ubx/lacZ* fusion. In the upper line, which represents the genomic DNA of the *Ubx* unit, the four boxed regions denote exon sequences retained in the spliced mRNA. The mRNA contains a 380-codon ORF (Ia, in the stippled region), of which codons 9-300 were incorporated into the *Ubx/lacZ* fusion (bottom line). The locations of the *Ubx* pseudo-point mutations (ref. 20; M. Akam, personal communication; W. Bender, unpublished results; P.A.B. and D.S.H., unpublished) are indicated above the upper line: *Ubx*^{6.28}, *Ubx*^{8.49} and *Ubx*^{9.22} are associated, respectively, with deletions of ~32 bp, 110 bp and 1.6 kb; *Ubx*¹ is associated with the insertion of a 'Doc' transposable element. The thin boxes indicate the positions of the homoeo box¹¹⁻¹⁴, and of the 'hinge region', which has been so named because of the many glycine codons. (Because of its high glycine content, the hinge region almost certainly exists within native *Ubx* proteins as a region devoid of secondary structure³⁴⁻³⁶, perhaps functioning as a flexible connection between spatially distinct domains.) The glycine content of the entire hinge region is 73% (22 of 30 amino acid residues) and higher (89%, 17 of 19 amino acid residues) in the filled portion of the box. The gene fusion was constructed by standard methods³⁷, using the λ gt11 expression vector³⁸. The *EcoRI* linker (BRL octamer) were selected to allow joining of the amino- and carboxy-terminal ends of the ORF Ia fragment (from *Ubx* cDNA p ϕ 3602) in-frame with *lacZ* gene sequences. A recombinant phage, λ Ldm3751, was identified by its *lacZ*⁻ phenotype, by hybridization to *Ubx* sequences and by production of a novel polypeptide of expected electrophoretic mobility in SDS gels on infection of *E. coli* host strain Y1089 (ref. 38; see Fig. 3).

Wilde and M. Akam, personal communication). In addition, Fig. 2 shows that of the four *Ubx* pseudo-point mutations which have been mapped, two (*Ubx*^{6,28}, *Ubx*⁸⁴⁹) are short deletions in the 5'-exon region common to both ORFs; one (*Ubx*^{9,22}) deletes the entire 3'-exon region of each, and one (*Ubx*¹) consists of an insertion in or near their 5'-exon regions. The 12 other *Ubx* loci shown in Fig. 1 represent breakpoints of gross chromosomal rearrangements that separate the ends of the *Ubx* unit by large genomic distances; although these breakpoints are less informative than the pseudo-point mutations, they are consistent with a coding function for the ORFs because they should prevent formation of the appropriate mRNAs by interrupting transcription of the unit.

Working model for the *Ubx* domain

The location of the pseudo-point deletions and the conclusion that *Ubx* mutations inactivate all functions of the domain, suggest that each of these functions requires one or more proteins encoded by the *Ubx* unit and that these *Ubx* proteins form a family characterized by common amino-acid sequences in their amino- and carboxy-proximal regions. The members of the family would then be distinguished by differences in the amino-acid sequence connecting these two constant regions, exemplified by the two proteins encoded by ORF Ia and b. Other members of the family may be encoded by the 4.3-kb RNAs, which also contain sequences from both ends of the *Ubx* unit and exhibit the same 5' initiation site as the 3.2-kb RNAs (R. B. Saint, P.A.B., D. Peattie, P. Harte and D.S.H., in preparation). Thus, we suppose that differential processing of the long *Ubx* primary transcript generates a set of mRNAs that encode the *Ubx* proteins and that specific subsets of these mRNAs are produced and translated in different compartments to provide the individual, compartment-specific functions of the domain. Certain predictions of this model are confirmed by the segmental distribution of the *Ubx* proteins described here.

Antibodies to *Ubx* proteins

The distributions examined here are those registered with a polyclonal antibody probe that should detect all members of the *Ubx* protein family because the antigen used to elicit these antibodies contains their common amino-proximal region. This antigen is a *Ubx*- β -galactosidase hybrid protein (fla1) synthesized in *Escherichia coli* from a gene fusion consisting of codons 9-300 of ORF Ia inserted in-frame into the *E. coli lacZ* gene of the λ gt11 expression vector, as indicated at the bottom of Fig. 2 and described in the Fig. 2 legend. To enrich for antibodies directed against *Ubx* determinants, the IgG fraction of serum from rabbits injected with purified fla1 was affinity-purified by binding to fla1, after which antibodies directed against β -galactosidase and other *E. coli* proteins were removed by adsorption (see Fig. 3 legend for details). Figure 3 shows that this antibody preparation reacts only with the *Ubx* determinants of fla1 and of another *Ubx*- β -galactosidase hybrid (f1b1) containing the amino acid sequence encoded by codons 9-389 of ORF Ib.

Segmental distributions of *Ubx* proteins

We have used indirect immunofluorescence staining with these anti-*Ubx* antibodies to examine the prediction of the model that at least one member of the *Ubx* protein family will be expressed in each compartment whose identity depends on the functions of the *Ubx* domain. These functions are defined by the general rule that mutational inactivation of a given function transforms the compartment in which it is normally expressed to the identity of the homologous compartment in the anteriorly adjacent segment, where it is not normally expressed. Four functions are required to explain the compartmental transformations in the larval and adult cuticle caused by the five classes of mutations in the *Ubx* domain. These functions are referred to here as ABX, BX, PBX and BXD and are expressed, respectively, in the T2p, T3a, T3p and Ala compartments. The transformations caused by each mutational class can be explained if the following functional inactivations are assumed: *Ubx* mutations inactivate

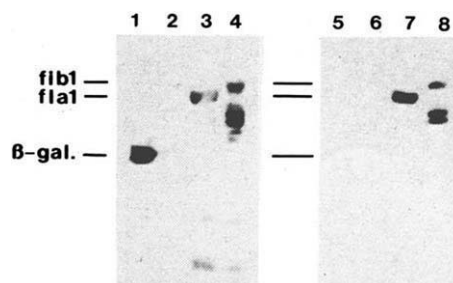


Fig. 3 Antibodies to *Ubx* determinants. A purified immunoglobulin fraction derived from hybrid protein-elicited antibodies reacts exclusively with *Ubx* determinants on SDS gel blots^{39,40} of total protein extracts from *E. coli* lysogens. This fraction was prepared in two steps: (1) by affinity purification of antibody using the hybrid protein as ligand; (2) subsequent removal of β -galactosidase-reactive material (lanes 1-4, specificity of antibody binding after step 1; lanes 5-8, specificity after steps 1 and 2). Thus, in lanes 7 and 8, the hybrid proteins and their cellular breakdown products react, whereas in lanes 5 and 6 neither β -galactosidase nor other *E. coli* proteins are bound. Protein extracts in lanes 2 and 6 were from the *lacZ*⁻ host cell, Y1089 (ref. 38). The remaining extracts were from induced lysogens of the following phage in the Y1089 host: lanes 1 and 5, the vector λ gt11, carrying the *lacZ*⁺ gene; lanes 3 and 7, λ Ldm3751; lanes 4 and 8, λ Ldm3752, a related construction incorporating codons 9-389 from ORF Ib (P.A.B. and D.S.H., unpublished data). The hybrid proteins encoded by λ Ldm3751 and λ Ldm3752 are labelled f1a1 and f1b1, respectively; their apparent M_r s are as predicted from the cDNA sequences (f1a1, 145,000; f1b1, 153,000).

Methods. After growth to $A_{600} = 0.5$ in L-broth, a lysogen of λ Ldm3751 (see Fig. 2) in Y1089 (ref. 38) was induced by a temperature shift (43°, 15 min) and addition of isopropyl thiogalactoside (IPTG; Sigma) to 5 mM. Cells were collected 45 min after IPTG induction and lysed by a freeze-thaw cycle following lysozyme digestion⁴¹. To the cleared lysate (supernatant of a 30 min, 200,000g centrifugation), solid ammonium sulphate was added to 35% of saturation at 0°C; the resulting precipitate was resuspended and dialysed (Fraction I) and loaded in sample application buffer on 6% SDS gels³⁹. After light staining with Coomassie Brilliant Blue, the band corresponding to the *Ubx*- β -galactosidase hybrid protein was excised, homogenized and emulsified with Freund's complete adjuvant. Primary immunization of virgin female New Zealand White rabbits was by subcutaneous (s.c.) injection of 250 μ g (estimated by comparison with protein standards in Coomassie-stained gels) of hybrid protein per rabbit at ~20 sites. Booster immunizations were given 67 days after the primary immunizations by s.c. injection at ~10 sites with 100 μ g of hybrid protein emulsified in Freund's incomplete adjuvant per rabbit. An IgG fraction was purified from serum collected 10 days after boosting⁴² and was purified further using an affinity adsorbent made by immobilizing Fraction I proteins (the hybrid protein constitutes 10-20% of the total protein in Fraction I) on derivatized beads (Affigel 10; Biorad)^{43,44}. Antibodies directed against β -galactosidase or other *E. coli* proteins were removed by exhaustive treatment with an adsorbent similarly made with a protein fraction from a lysogen of λ gt11 bearing no insert.

all functions; *abx* and *bxd* mutations inactivate the ABX, BX and BXD, PBX pairs of functions, respectively; *bx* and *pbx* mutations inactivate the BX and PBX functions, respectively (refs 1, 2, 4-10; G. Morata, personal communication).

According to the model, we therefore expect that *Ubx* proteins will be found in the cell lineages for T2p, T3a, T3p and Ala and will not be present in more anterior compartments, which are not affected by the five classes of mutation. Posterior to Ala, however, the situation is more complex. Although the primary BX-C determinants for segmental identities in the remainder of the abdomen derive from the other two BX-C domains^{1,3}, *Ubx* and *bxd* mutations cause subtle transformations of the larval cuticle in the anterior portions of A2-7, indicating that a BXD-related function is expressed in the presumed anterior compartments of these segments¹. Thus, the model predicts that the cells of the anterior compartment lineages will also contain *Ubx* proteins.

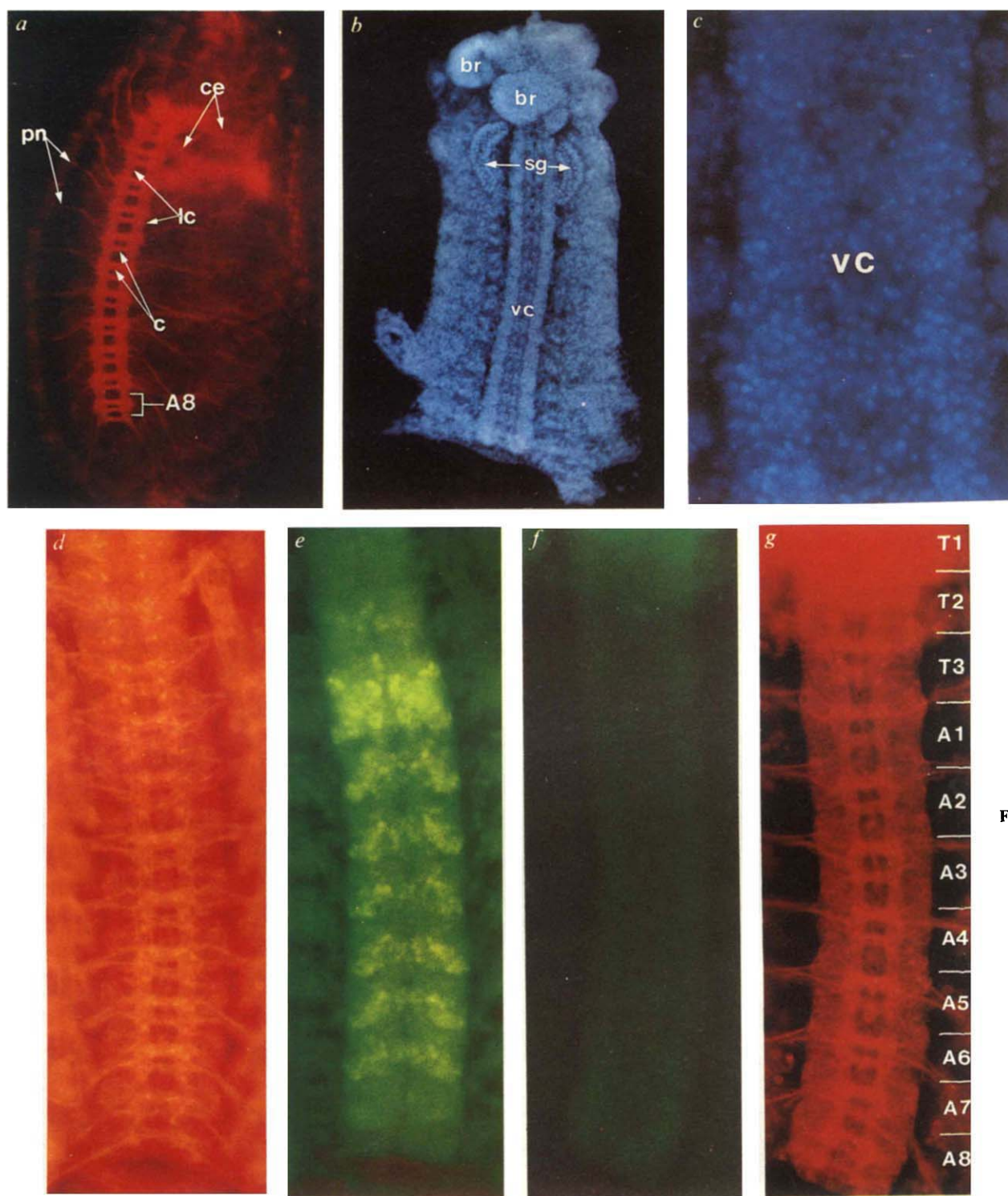


Fig. 4

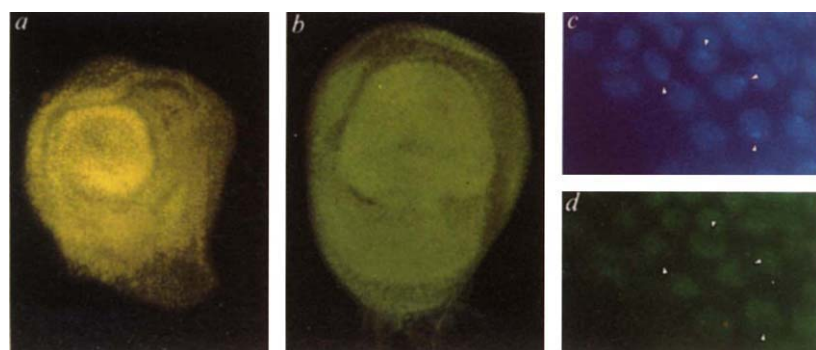


Fig. 5

We have examined the distribution of *Ubx* proteins in the central nervous system (CNS) at mid-embryogenesis and in the imaginal disks that are the precursors of the adult cuticular segments. The embryonic CNS is excellent for this purpose because it is easily accessible, expresses *Ubx* transcripts at relatively high concentrations¹⁵ and its ventral nerve cord consists of readily distinguishable paired ganglia corresponding to T1–A8, as well as the more anterior pair of suboesophageal ganglia and the incomplete ninth abdominal ganglia (Fig. 4a–c, g). Figure 4d, e shows the simultaneous staining of a wild-type CNS with anti-tubulin and anti-*Ubx* antibodies, respectively. The anti-tubulin staining reveals the ladder-like axon scaffolding overlying the cell bodies of the nerve cord, whose paired rungs (commissures) provide the landmarks for identifying the segmental boundaries¹⁶. Using these landmarks, weak anti-*Ubx*

staining is observed in the posterior portion of T2, followed by strong staining in posterior T3 and anterior A1, with the anterior portions of A2–7 exhibiting progressively weaker staining (see Fig. 6). Except for the very weak staining of anterior T3, the qualitative aspect of this pattern is that expected from the model, provided that the anti-*Ubx* staining is specific to *Ubx* proteins.

Figure 4f, g removes this proviso by showing that the anti-*Ubx* staining disappears in the embryonic CNS from a *Ubx*^{6,28} homozygote. Preliminary sequence analysis (M. Akam, personal communication) indicates that the small *Ubx*^{6,28} deletion near the 5' end of ORF 1a and b (Fig. 2) shifts the reading frame to yield a stop codon downstream from the deletion. Clearly, the absence of protein interacting with the anti-*Ubx* antibodies is consistent with this sequence analysis and demonstrates that these antibodies react uniquely with the *Ubx* family of proteins.

Figure 5a, b shows that the haltere and third leg imaginal disks that form the T3 segment of the adult cuticle exhibit strong anti-*Ubx* staining. The wing and second leg disks of adult T2 exhibit little or no anti-*Ubx* staining, whereas the T1 first leg disk and the eye-antennal and genital disks are not stained significantly (data not shown). The results for the haltere and wing disks are consistent with the observation that the haltere (but not the wing) disk contains both the 3.2-kb and 4.3-kb size classes of *Ubx* RNAs (R. B. Saint, P.A.B., D. Peattie, P. Harte and D.S.H., in preparation). Furthermore, the pattern of anti-*Ubx* staining of both disks and the mid-embryonic CNS overlaps precisely the distribution of RNAs homologous to the 5' exon of the *Ubx* unit, as registered by *in situ* hybridization¹⁵, although the level of resolution for the RNA distributions is reduced considerably.

The sum of the CNS and disk distributions for *Ubx* proteins clearly fits the spatial localizations predicted by the model. Qualitative differences between the distributions are restricted to posterior T2 and anterior T3. The posterior T2 difference can be explained simply if it is assumed that ABX expression is required at the same time period in development of the adult cuticle and the CNS. Clonal analyses indicate that normal T2p development of the adult cuticle requires ABX function only during the first ~8 h of embryogenesis (ref. 6 and G. Morata, personal communication). As the CNS was assayed only a few hours after this critical period and the disks were assayed ~100 h later, the staining of posterior T2 in the embryonic CNS, but not the disks, may simply reflect the instability or dilution of the *Ubx* proteins and mRNAs associated with the ABX function.

The difference in staining of anterior T3 in the CNS and disks is of greater interest because it includes a staining difference between the two T3 disks. Haltere disk cells are stained strongly and uniformly with no anterior-posterior distinction. By contrast, cells of the third leg disk stain with markedly different intensities in defined regions, the most striking difference resulting from the weaker staining anterior compared with posterior cells. Thus, the anterior T3 staining increases from slight in the CNS to moderate and strong, respectively, in the third leg and haltere disks. This increase is correlated with the progressively more dorsal position of the precursors of these tissues on the dorsal-ventral axis of the blastoderm fate map¹⁷, suggesting that the amount and/or combination of *Ubx* proteins expressed may be regulated according to position on this axis, just as it is regulated according to position on the anterior-posterior axis.

In this regard, it is of interest that *abx* mutations affect the development of posterior T2 and anterior T3 in both the CNS and the cuticle (ref. 18 and G. Morata, personal communication), whereas *bx* mutations produce little if any detectable effect in the CNS¹⁸. Thus, the dependence of BX function on *bx*⁺ sequences observed in adult cuticle development does not obtain in the CNS, perhaps because of regulation along the dorsal-ventral axis such as that mentioned in the previous paragraph. By contrast, the ABX function defined for cuticle development could exist without modification in the CNS, as the requirement for *abx*⁺ sequences is common to both.

Fig. 4 (left) *Ubx* proteins in the CNS at mid-embryogenesis. a–c, Fluorescence photomicrographs showing the neuropil (a) and cell nuclei (b, c) of the 12-h CNS. d–g, Simultaneous staining with anti-tubulin and anti-*Ubx* antibodies of CNS dissections from Oregon-R wild-type and *Ubx* mutant embryos. The axon bundles of the neuropil are visualized by indirect immunofluorescence staining (a) of intact embryos⁴⁵ using an anti-tubulin antibody (supplied by T. Karr from a hybridoma cell line established by S. Blose). In b and c, the CNS has been exposed by dissection¹⁶, and the nuclei are visualized with Hoechst 22358 (bisbenzidine), a fluorescent DNA-binding dye⁴⁵. In c, individual nuclei can be distinguished as brightly fluorescing points in a field of more diffuse light, which comes from the nuclei above and below the plane of focus. The paired commissures of the neuropil can be used as landmarks to identify the ganglia of the ventral cord: three thoracic pairs and eight complete abdominal pairs (the ninth abdominal ganglia are incompletely formed⁴⁶; E. Teugels and A. Ghysen, personal communication). In embryos at this stage of development, a slight longitudinal compression of the commissural ladder is apparent in A8, perhaps due to the onset of ventral cord condensation⁴⁷. Abbreviations: br, brain; ce, circumoesophageal connectives; c, commissures; lc, longitudinal connectives; pn, peripheral nerves; sg, salivary glands; vc, ventral cord. Oregon-R embryos displayed consistently a characteristic staining pattern with both antibodies (d, e) and the anti-*Ubx* stain (e) proved to be nuclear on comparison with bisbenzidine staining patterns (see Fig. 6 for a schematic representation). Of the 38 embryos from the balanced heterozygous stock examined (*Ubx*^{6,28}/TMI; see Fig. 2), 10 (26%) did not stain above background with anti-*Ubx* antibodies (f), although the neuropil (g) appeared normal (colour differences of f and g from e and d are due to the use of a different photographic emulsion). CNS dissections and antibody staining were as described in ref. 16, and secondary antibodies were from Sigma (fluorescein-conjugated goat anti-rabbit IgG) and Cappel (rhodamine-conjugated goat anti-mouse IgG). The anti-*Ubx* antibodies used for CNS experiments were prepared as for lanes 5–8 in Fig. 3, but in addition were treated to remove any antibodies cross-reacting with other *Drosophila* antigens. The adsorbent for this treatment consisted of de-vitelinized embryos⁴⁵ which had been fixed at a stage before 2 h after fertilization, the time at which *Ubx* RNAs are first expressed.

Fig. 5 (left) *Ubx* proteins in imaginal disks. a, b, Anti-*Ubx* staining of the T3 imaginal disks; c, d, nuclear localization of *Ubx* proteins. Anti-*Ubx* antibodies react strongly with cells of the haltere disk and the third leg disks of third-instar larvae (a and b respectively; fluorescence photomicrographs were taken under slightly different filter conditions), but only weakly, if at all, with cells of other imaginal disks. Whereas cells of the haltere disk stain uniformly, cells of the third leg disk stain with markedly different intensities in defined regions. The most striking difference results from the weaker staining of anterior compared with posterior cells, but intensity differences between proximal and distal cells are also observed. Anti-*Ubx* staining (d) can be seen to co-localize with bisbenzidine fluorescence (c) in a high-magnification view of a field of cells from a partially everted haltere disk. *Ubx* proteins are apparently excluded from a region of the nucleus, possibly the chromocentre, which stains very brightly with bisbenzidine (see arrowheads). Disks were stained as described in Zipursky *et al.*⁴⁸. All photomicrographs in this figure Fig. 4 were taken under epifluorescence illumination with Zeiss optics.

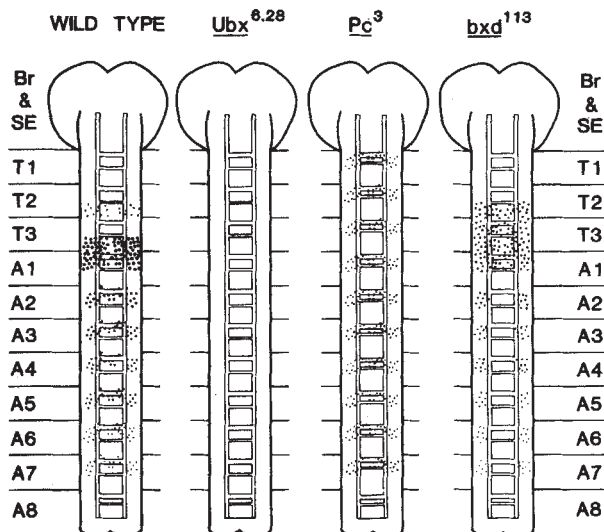


Fig. 6 Spatial distribution of *Ubx* proteins in the CNS of wild-type and mutant embryos. The results of anti-*Ubx* staining of CNS dissections of wild-type and mutant embryos are shown in schematic form. The number of dots in a given area is proportional to the number of nuclei stained and the relative staining intensities are indicated by dot size. In the wild-type CNS the most brightly staining nuclei are located in a region comprising posterior cells of the T3 ganglia (T3p) and anterior cells of the A1 ganglia (A1a). Most cells in this region are stained. More anteriorly, a subset of the cells in T2p is stained lightly, while the staining of T3a is so faint that it is not shown in the diagram. Posterior to the brightly staining T3p-A1a region, stained cells in segments A2-7 form a subset of the total complement of anterior cells. The number and arrangement of nuclei in these stained anterior patches is similar if not identical from segment to segment, although the intensity of staining decreases progressively in the posterior direction. (The extremely faint staining observed in A8 is not depicted.) Staining patterns of embryos collected from balanced heterozygous mutant stocks fell into two classes, of which the larger was indistinguishable from wild type and the smaller displayed the altered pattern shown in the figure. This variant class is presumed to represent the mutant homozygote as it generally comprised about a quarter of the embryos examined. Of the embryos collected from the *Ubx*^{6.28}/TM1 stock, 26% (10 of 38 individuals examined) did not stain above background levels, although the neuropil displayed no evident differences from the wild type (Fig. 4g). The variant classes from both *Pc* stocks examined (*Pc*³/TM3 and *Pc*²/TM1)^{22,49} displayed similar differences in *Ubx* protein distribution and neuropil structure. These differences could be summarized as the pre-emption of the normal features of the thoracic and anterior abdominal segments by the features normally seen only in the most posterior abdominal segments. The proportion of abnormal individuals in the combined progeny from both *Pc* stocks was 28% (17 of 61). The variant class of the *bxd*¹¹³ stock examined (In(3)*bxd*¹¹³/TM1)¹¹ displayed the modified staining pattern shown above, with the neuropil retaining its normal features; 10 individuals of the 31 examined (32%) from the *bxd*¹¹³ stock showed the abnormal pattern. Collection and staining of embryos was as described in Fig. 4 legend.

Regulation of the *Ubx* unit

One consequence of our model is that the *bxd* unit acts as a *cis*-regulator of the *Ubx* unit. This follows from the assumption that the individual functions of the *Ubx* unit, including those that are *cis*-inactivated by *bxd* mutations, are executed by specific combinations of *Ubx* proteins. The *bxd* unit delineated by these mutations would then *cis*-regulate the formation of particular members of the *Ubx* protein family, presumably by facilitating and/or inhibiting particular pathways by which the respective mRNAs are processed from the *Ubx* primary transcript. The idea that the *bxd* unit regulates the *Ubx* unit is also suggested by recent genetic evidence (ref. 19 and E. B. Lewis, personal communication).

Teugels and Ghysen¹⁸ showed that extreme *bxd* mutations alter CNS development of posterior T3 and anterior A1, consistent with their effects on cuticular development. The strong anti-*Ubx* staining of these two regions (Figs 4e, 6) may then result from the combinations of *Ubx* proteins that execute the functions inactivated by these *bxd* mutations (that is, BXD and PBX). The staining of the anterior portions of A2-7 in the CNS may also be indicative of *Ubx* proteins with BXD-related functions, although this suggestion derives solely from observations of the subtle effects of *bxd* mutations on these segments in the larval cuticle¹. To test these predictions, we have examined the distribution of *Ubx* proteins in the mid-embryonic CNS of the extreme *bxd*¹¹³ mutation (an inversion associated with a deletion of DNA at the breakpoint in the *bxd* unit) and of the weaker *bxd*¹ and *bxd*⁵⁵¹ mutations, each of which results from the insertion of a 'gypsy' transposable element in the *bxd* unit²⁰.

No obvious change in the *Ubx* protein distribution was observed in homozygotes for the weak *bxd* mutations, which is not surprising as Teugels and Ghysen observe only minor alterations in the CNS in a *bxd*¹ hemizygote¹⁸. By contrast, Fig. 6 shows that the extreme *bxd*¹¹³ mutation alters dramatically the segmental distribution of *Ubx* proteins. The expected regions are altered; the staining intensity of posterior T3 and anterior A1 cells is reduced markedly from that in wild type, as is the intensity in the anterior regions of A2-7, where the number of stained cells is also reduced by about twofold in a spatially specific manner. Both intensity and number of stained cells, however, is increased in posterior T2 and anterior T3; indeed, these two regions together with posterior T3 and anterior A1 form a block in which most cells are stained equivalently, as if the *bxd*¹¹³ mutation has relaxed a fine control over *Ubx* protein formation at a compartment or lower level. The change of pattern in posterior T2 and anterior T3 is not entirely unexpected, as in these regions, *bxd* mutations are known to modify the effects of other *Ubx*-related mutations (ref. 21 and E. B. Lewis, personal communication); in addition, changes in what is thought to be T3a of the CNS were observed by Teugels and Ghysen for the *bxd* mutations they examined¹⁸. Whatever the morphological transformations, the striking change in *Ubx* protein distribution caused by *bxd*¹¹³ strongly suggests that the *bxd* unit regulates expression of the *Ubx* unit.

Trans-regulation of *Ubx* protein distribution

The *Polycomb* gene (*Pc*) is a *trans*-regulator of BX-C expression located in the left arm of chromosome 3, far from the BX-C in the right arm²². Embryos lacking *Pc*⁺ function suffer a massive homoeotic transformation by which all thoracic and abdominal cuticle segments exhibit identities characteristic of normal A8. Based on genetic evidence that this transformation requires the presence of the BX-C in the genome, Lewis¹ postulated that the *Pc*⁺ product acts as a negative regulator of the expression of all genes in the BX-C. All BX-C genes were included in the scope of this regulator because he had proposed on other grounds that all BX-C genes are expressed in A8 (ref. 1); thus, it was argued that the transformation of all segments to A8 was equivalent to the activation of all BX-C genes in all segments by inactivation of the *Pc*⁺ negative regulator.

We have tested this hypothesis by determining the distribution of *Ubx* proteins in the CNS of *Pc*² and *Pc*³ homozygous embryos. Figure 6 shows that both the neuropil structure and the *Ubx* protein distribution are transformed in a manner closely approximating the cuticle transformation: the spacing between the paired commissures for the T1-A7 segments is reduced to the lesser spacing of wild-type A8 and the anti-*Ubx* staining pattern of segments T1-A7 appears intermediate between wild-type A7 and A8 (wild-type A8 staining is so faint that it is not noted in Fig. 6), although there is a gradual increase in intensity anteriorly. These results, however, are inconsistent with the simple hypothesis of a single *Pc* regulator for the BX-C. Clearly, other genes than *Pc* must function as regulators of the *Ubx* domain to account for the distribution observed in *Pc* mutants; indeed, there is a growing list of putative *trans*-regulators of

the BX-C^{19,23-28} that may exhibit such a function and can be tested easily in the above manner. In addition, we are testing the possibility that other genes of the BX-C function in the abdomen to regulate expression of the *Ubx* domain. If these genes were themselves regulated by the *Pc*⁺ product, they would be expressed ubiquitously in *Pc* mutants, resulting in the observed repetitive abdominal-like pattern of *Ubx* proteins.

Ubx proteins are nuclear proteins

Figure 5c, d shows that *Ubx* proteins are localized in the nuclei of the haltere disk cells. Nuclear localization was also observed for third leg disk and embryonic CNS cells and seems to be a general property of the *Ubx* proteins relevant both to hierarchical models of development (in which the BX-C products are postulated to directly regulate other genes^{29,30}) and to speculation that homoeo boxes endow proteins with sequence-specific DNA-binding capabilities^{31,32}.

The homoeo box of the *Ubx* unit is located just downstream from the acceptor splice site for the 3' exon region common to ORF Ia and b (Fig. 2). In our model for the *Ubx* protein family, the homoeo box is located in the carboxy-proximal constant region, near its boundary with the variable region. This location allows juxtaposition of the homoeo box with a variety of amino acid sequences that could modulate its function to create a diversity of related regulatory proteins. We emphasize that the function of the homoeo box need not be to bind DNA. For example, one can imagine equally that the variable region provides that function while the homoeo box provides a site by which *Ubx* proteins bind to a conserved protein, such as a histone, to create a perturbation in chromatin structure that has negative or positive regulatory value. Indeed, the *Ubx* proteins may regulate RNA processing rather than DNA transcription; it is known only that they accumulate in nuclei of segments whose identities are regulated by the *Ubx* domain.

Conclusions

The primary assumption of our working model for the *Ubx* domain is that all functions defined by mutations in the domain are executed by the proteins encoded by the *Ubx* transcription unit. That assumption leads to the prediction, confirmed qualitatively here, that the sum of the segmental distributions for these *Ubx* proteins will follow the sum of the segmental distributions for the functions of the domain. A second prediction of our model is that mutations in the other transcription unit of the domain, the *bxd* unit, will change the segmental distribution of *Ubx* proteins. This prediction has also been confirmed in a

preliminary fashion; other *bxd* mutations, including their *pbx* relatives, must be tested, as must the prediction that this *bxd* effect is limited to *cis*-interactions.

A second assumption of the model is that a family of related proteins is encoded by the *Ubx* unit and that different combinations of these *Ubx* proteins execute the different functions of the domain, suggested by the observation that the *Ubx* unit produces a set of putative mRNAs, two of which contain ORFs that differ by a single step in their splicing pathways (ORF Ia and b). We have begun to test this part of the model by two methods: (1) the further definition of the *Ubx* RNAs via cDNA sequence analysis and the use of that sequence data to generate antibody probes specific to the different ORFs they may contain. For example, we have synthesized a peptide containing the nine-amino acid sequence that distinguishes ORF Ia from b and raised antibody against it for the identification of the ORF Ib protein. (2) We are performing a characterization of the proteins in *D. melanogaster* that react with the anti-*Ubx* antibodies. Although we have found multiple proteins in embryonic nuclear extracts with appropriate mobilities in SDS gels, the source of the apparent variation in *M_r* has not yet been determined. Similar observations were reported recently by White and Wilcox³³, who also obtained results similar to those reported here on the anti-*Ubx* staining of wild-type imaginal disks.

Perhaps the most novel characteristic of the *Ubx* domain is that its *cis*-interacting functions depend on two non-overlapping, commonly oriented transcription units; the most novel aspect of our model is that we assign the protein-coding functions of the domain to the unit containing the homoeo box and ascribe *cis*-regulatory functions to the other. The possibility that this novel organization occurs in other homoeotic loci is suggested by recent evidence that each of the other two domains of the BX-C contains a single homoeo box and multiple transcription units (F. Karch and W. Bender, personal communication; S. Sakonju and D.S.H., unpublished experiments).

We thank Corey Goodman for advice and equipment, Tim Karr for advice and antibodies and Michael Akam, Deborah Wilde, Welcome Bender, Francois Karch, Alain Ghysen, Erik Teugels, Ed Lewis and Gines Morata for communication of unpublished results. We also thank members of the Stanford Department of Biochemistry for advice and help, particularly, Howard Lipshitz and Renato Paro for mutant flies and Robert Baldwin and Peter Kim for discussions about protein structure. We also thank Jeremy Nathans and Howard Lipshitz for comments on the manuscript. This work was supported by a NIH grant to D.S.H.; P.A.B. was supported by a NSF graduate fellowship and a NIH training grant and S.L.H. by a NIH postdoctoral fellowship.

Received 31 December 1984; accepted 17 January 1985.

- Lewis, E. B. *Nature* **276**, 565-570 (1978).
- Lawrence, P. A. & Morata, G. *Cell* **35**, 595-601 (1983).
- Sanchez-Herrero, E., Vernos, I., Marco, R. & Morata, G. *Nature* **313**, 108-113 (1985).
- Lewis, E. B. *Am. Zool.* **3**, 33-56 (1963).
- Lewis, E. B. in *Developmental Biology Using Purified Genes* (eds Brown, D. D. & Fox, C. F.) 189-208 (Academic, New York, 1981).
- Morata, G. & Kerridge, S. *Nature* **290**, 778-781 (1981).
- Kerridge, S. & Morata, G. *J. Embryol. exp. Morph.* **68**, 211-234 (1982).
- Minana, F. J. & Garcia-Bellido, A. *Wilhelm Roux Arch. dev. Biol.* **190**, 331-334 (1982).
- Hayes, P. H., Sato, T. & Denell, R. E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 545-549 (1984).
- Struhl, G. *Nature* **308**, 454-457 (1984).
- McGinnis, W. *et al. Nature* **308**, 428-433 (1984).
- Scott, M. P. & Weiner, A. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4115-4119 (1984).
- McGinnis, W. *et al. Cell* **37**, 403-408 (1984).
- Carrasco, A. E. *et al. Cell* **37**, 409-414 (1984).
- Akam, M. *EMBO J.* **2**, 2075-2084 (1983).
- Thomas, J. B., Bastiani, M. J., Bate, M. & Goodman, C. S. *Nature* **310**, 203-207 (1984).
- Garcia-Bellido, A. & Ripoll, P. in *Genetic Mosaics and Cell Differentiation* (ed. Gehring, W. J.) 119-156 (Springer, Berlin, 1979).
- Teugels, E. & Ghysen, A. *Nature* (submitted).
- Ingham, P. W. *Cell* **37**, 815-823 (1984).
- Bender, W. *et al. Science* **221**, 23-29 (1983).
- Lewis, E. B. *Am. Nat.* **89**, 73-89 (1955).
- Lindsley, D. L. & Grell, E. H. *Genetic Variations of Drosophila melanogaster* (Carnegie Instn Wash. Publ. 627, 1969).
- Ingham, P. & Whittle, R. *Molec. gen. Genet.* **179**, 607-614 (1980).
- Capdevila, M. P. & Garcia-Bellido, A. *Wilhelm Roux Arch. dev. Biol.* **190**, 339-350 (1981).
- Struhl, G. *Nature* **293**, 36-41 (1981).
- Duncan, I. *Genetics* **102**, 49-70 (1982).
- Duncan, I. & Lewis, E. B. in *Developmental Order: Its Origin and Regulation* (ed. Subtelny, S.) 533-554 (Liss, New York, 1982).
- Ingham, P. W. *Nature* **306**, 591-593 (1983).
- Garcia-Bellido, A. *Ciba Fdn Symp.* **29**, 161-182 (1975).
- Garcia-Bellido, A. *Am. Zool.* **17**, 613-629 (1977).
- Laughon, A. & Scott, M. P. *Nature* **310**, 25-31 (1984).
- Shepherd, J. C. W. *et al. Nature* **310**, 70-71 (1984).
- White, R. A. H. & Wilcox, M. *Cell* **39**, 163-171 (1984).
- Ananthanarayanan, V. S., Andreatta, R. H., Poland, D. & Scheraga, H. A. *Macromolecules* **4**, 417-424 (1971).
- Scheraga, H. A. *Pure appl. Chem.* **50**, 315-324 (1978).
- Schulz, G. E. & Schirmer, R. H. *Principles of Protein Structure* (Springer, New York, 1979).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Young, R. A. & Davis, R. W. *Science* **222**, 778-782 (1984).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Burnette, W. N. *Analyt. Biochem.* **112**, 195-203 (1981).
- Fuller, R. S., Kaguni, J. M. & Kornberg, A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7370-7374 (1981).
- Van der Zeijst, B. *et al. J. Virol.* **48**, 249-261 (1983).
- Ikeda, Y. & Steiner, M. *J. biol. Chem.* **251**, 6135-6141 (1976).
- Kristiansen, I. in *Affinity Chromatography* (ed. Hoffman-Ostenhoff, O.) 191-206 (Pergamon, Oxford, 1978).
- Mitchison, T. J. & Sedat, J. *Dev. Biol.* **99**, 361-264 (1983).
- Hartenstein, V. & Campos-Ortega, J. A. *Wilhelm Roux Arch. dev. Biol.* **193**, 308-325 (1984).
- Kankel, D. R. *et al. in The Genetics and Biology of Drosophila* (eds Ashburner, M. & Wright, T. R. F.) 295-369 (Academic, London, 1980).
- Zipursky, S. L., Venkatesh, T. R., Teplow, D. B. & Benzer, S. *Cell* **36**, 15-26 (1984).
- Lewis, E. B. *Dros. Inf. Serv.* **55**, 207-208 (1980).

Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors

Maria M. Konarska*, Paula J. Grabowski, Richard A. Padgett & Phillip A. Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The branch site of lariat RNAs produced during the splicing of the first two late leader exons of adenovirus-2 has a structure of ... A₃^{2p5'G...}U... There is a distinct complementarity between sequences preceding the adenosine at the branch site and the 5' terminus of the intervening sequence.

EUKARYOTIC genes frequently contain intervening sequences that are removed from precursor RNAs by splicing. Introns occur in many genes coding for proteins, transfer RNAs and ribosomal RNAs in nuclear, mitochondrial and chloroplast DNAs¹⁻³, yet no single general mechanism seems to be involved in this splicing. Detailed studies of *in vitro* splicing of precursor tRNAs revealed that two ligation pathways operate in plant and animal systems⁴⁻⁷. The plant pathway involves ligation of two tRNA exons with 5'-phosphate and 2'-3'-cyclic phosphate ends,

leading to the formation of a N_{3'}^{2'}—P—5'N bond^{4,5}. The cyclic phosphate is displaced to the 2' position in the above structure. In precursor (pre)-tRNA splicing in HeLa cell extracts, however, two exons with a 5' hydroxyl group and 2',3'-cyclic phosphate ends are joined in a normal phosphodiester bond^{6,7}. A completely different mechanism operates in the splicing of rRNA

precursors in the protozoan *Tetrahymena*⁸. A series of transesterification reactions initiated by an attack of the 3' hydroxyl group of a guanosine cofactor is catalysed by the RNA molecule itself^{9,10}. The products consist of a normal phosphodiester bond linking the two exons and the excised intervening sequence that is subsequently circularized.

The recent developments of *in vitro* systems capable of splicing precursors of messenger RNA allows the mechanism of the reaction to be elucidated¹¹⁻¹⁵. Using a pre-RNA containing sequences from the adenovirus-2 (Ad-2) major late transcription unit and a whole-cell extract of HeLa cells, we have shown previously that splicing requires Mg²⁺ and ATP (ref. 14) and is inhibited by antisera against small nuclear ribonucleoprotein particles containing U1 RNA¹⁶. In this system, a 5'-terminal cap structure on the precursor RNA is required for splicing¹⁷. The reaction results in the formation of the spliced exons and the intervening sequence excised as a lariat RNA^{18,19}. Early in the reaction a probable intermediate, consisting of the lariat form of the intervening sequence joined to the 3' exon, was detected^{18,19}. We have shown previously¹⁹ that both lariat RNAs contain an unusual nuclease-resistant structure, an oligonucleo-

* Permanent address: Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02-532 Warsaw, Poland.

Fig. 1 *a*, Time course of the appearance in a HeLa nuclear extract of the products of splicing using the SP6 RNA polymerase-generated pre-RNA. Pre-RNA labelled with all four [α -³²P]NTPs was incubated in a HeLa cell extract with (lanes 1-5) or without (lanes 6-10) 1.5 mM ATP for the times indicated above the lanes. The RNA products were analysed on a 10% polyacrylamide 8 M urea gel; lane 11, ³²P-DNA markers from *Hinc*II digest of Φ X174 DNA. Left, structures of the various RNAs produced in the reaction; below, the structure of the DNA transcription template. L1 and L2 are the first and second leader exons of the Ad-2 major late transcription unit; IVS-1 is a shortened form of the first intervening sequence; IVS-2 a portion of the second intervening sequence; the hatched box represents sequences upstream of the *in vivo* RNA start site present in the RNA made with SP6 RNA polymerase. *b*, Analysis of the nuclease P₁-digestion products of purified IVS-A by two-dimensional TLC using solvent A in the first dimension and solvent B in the second dimension. *c*, The branched oligonucleotide was eluted from the plate in *b* and digested with snake venom phosphodiesterase. The products were analysed by two-dimensional TLC using solvent A in the first dimension and solvent C in the second dimension.

Methods. Pre-RNA (460 nucleotides) was transcribed from a *Bgl*II-cleaved pRSP-1- Δ IVS DNA template with SP6 RNA polymerase using a G(5')ppp(5')G dinucleotide primer¹⁷. Nuclear extract was prepared from uninfected HeLa cells as described previously²². Standard splicing reactions (25 μ l) contained: 50,000 c.p.m. ³²P-RNA precursor, 1.5 mM ATP, 5 mM creatine phosphate, 2.5 mM MgCl₂, 44 mM KCl, 8.8 mM HEPES pH 7.6, 0.2 mM EDTA, 0.9 mM dithiothreitol, 7.5% glycerol and 11 μ l nuclear extract. After 2 h of incubation at 30 °C, RNA was extracted and analysed on a 10% polyacrylamide 8 M urea gel¹⁸. Preparative reactions contained 5–25 $\times$ 10⁶ c.p.m. ³²P-RNA and were carried out in a total volume of 100 μ l. IVS-A RNA from a preparative splicing reaction was purified by polyacrylamide gel electrophoresis. Nuclease P₁ digestions were carried out in 10 μ l of 50 mM ammonium acetate pH 5.3 with 1 μ g enzyme for 1 h at 50 °C. Snake venom phosphodiesterase digestions were carried out in 10 μ l of 50 mM Tris-HCl pH 7.8 with 1 μ g enzyme for 1 h at 37 °C. Products of the digestions were separated by two-dimensional TLC on cellulose plates (Merck) using solvent A (isobutyric acid/concentrated NH₄OH:H₂O; 577:38:385), solvent B (*t*-butanol/concentrated HCl/H₂O; 14:3:3) or solvent C (saturated (NH₄)₂SO₄/1 M sodium acetate/isopropanol; 80:18:2). In the experiment described here, the ratio of incorporation of [α -³²P]-NTPs into pre-RNA was as follows: A/1.0, G/2.0, U/0.7, C/–1.5.

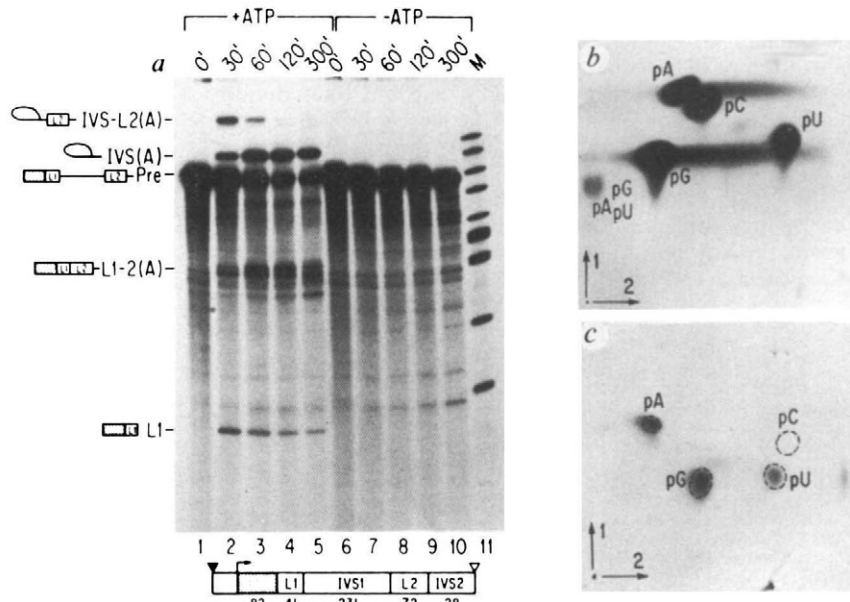
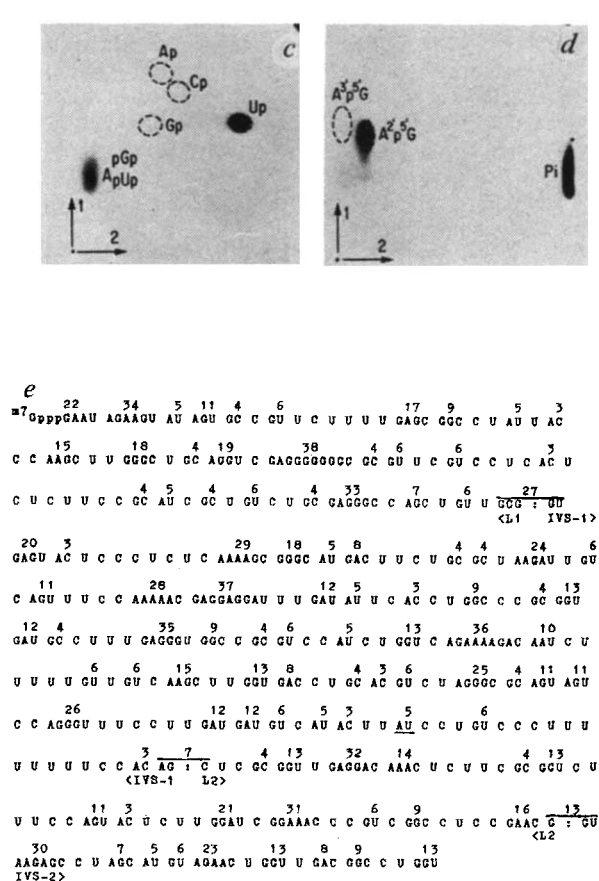
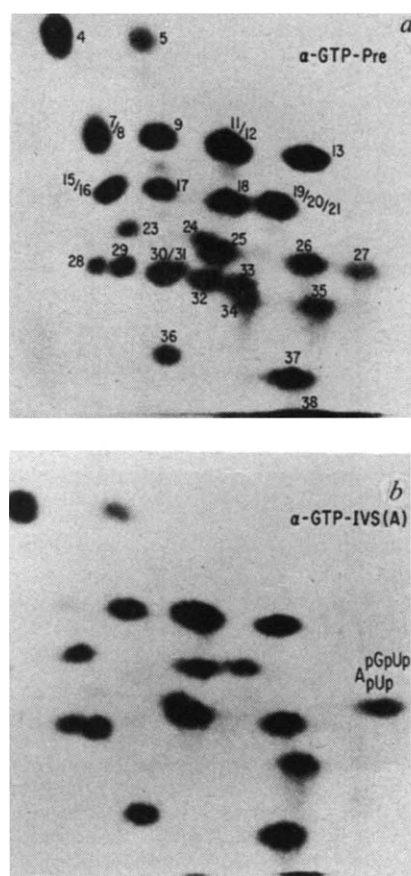


Fig. 2 *a*, Two-dimensional RNA fingerprint analysis of an RNase A digestion of purified pre-RNA labelled with [α - 32 P]GTP. *b*, Two-dimensional RNA fingerprint analysis of RNase A digestion of purified IVS-A RNA labelled with [α - 32 P]-GTP. *c*, The branch-containing oligonucleotide from *b* was digested with RNase T₁ and analysed by TLC in solvent A in the first dimension and solvent B in the second dimension. *d*, The branch-containing oligonucleotide from *b* was treated with phosphatase followed by periodate cleavage and β -elimination followed by phosphatase. The products were analysed by TLC as in *c*. The positions of the unlabelled A 2'p5'G and A 3'p5'G markers (Sigma) are indicated. *e*, Sequence of pre-RNA with all RNase A oligonucleotides numbered. Sequences corresponding to exons and introns are indicated and the AU dinucleotide at the branch site is underlined.

Methods. RNA fingerprint analysis. RNA was dissolved in 20 μ l 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and digested with 1.25 μ g RNase A for 2 h at 37 °C. The sample was lyophilized, resuspended in 3 μ l of H₂O and the oligonucleotides separated by electrophoresis on cellulose acetate from left to right followed by homochromatography on



DEAE-cellulose plates from bottom to top as described previously³². The branch-containing oligonucleotide was isolated from the DEAE-cellulose plate by elution with 1 M triethylamine bicarbonate, then digested with RNase T₁ (Calbiochem) (2 U in 50 mM Tris-HCl, pH 7.7) for 3 h at 37 °C or with calf alkaline phosphatase (Boehringer) (1 U in 50 mM Tris pH 7.7) for 1.5 h at 37 °C. Phosphatase-treated branched oligonucleotide was subjected to periodate oxidation in 10 μ l 0.1 M NaIO₄, 0.1 M sodium acetate pH 5.3 for 2 h at 20 °C in the dark followed directly by aniline treatment in 25 μ l of 0.5 M aniline 0.2 M sodium acetate, pH 5.3, for 3 h at 20 °C in the dark³³. The sample was extracted with phenol/chloroform (1:1), lyophilized and digested with phosphatase as above.

tide which is similar to the branch trinucleotide described previously²⁰, containing 2'-5' and 3'-5' phosphodiester bonds joined to a single adenosine residue¹⁹. The *in vitro* splicing of human β -globin pre-mRNA is similar²¹. In neither study was a complete analysis of a branch oligonucleotide reported.

We present here a detailed characterization of the branch oligonucleotide. Specifically we show that: (1) the branch trinucleotide has a structure of A₃^{2'p5'G}U; (2) the 5'-terminal G of the intervening sequence is joined directly to the adenosine residue at the branch site; (3) the phosphate group from the 5' splice site is incorporated into the 2'-5' phosphodiester bond; (4) the branch site is 24 nucleotides upstream of the 3' splice site; and (5) the probable intermediate in the reaction, the 5' exon, terminates in a 3' hydroxyl group.

Substrate and reaction conditions

As a model substrate in our studies, we have used a precursor RNA containing the first two leader exons (L1 and L2) of the Ad-2 major late transcription unit separated by a deleted form of the intervening sequence¹⁸. Pre-RNA was transcribed by SP6 RNA polymerase from a DNA template containing an SP6 promoter cloned upstream of the Ad-2 sequences¹⁷. Transcription of this substrate was primed using a G(5')ppp(5')G dinucleotide¹⁷ and in most cases, the pre-RNA processed by incubation in a HeLa nuclear extract²². All the characteristic products and intermediates described previously for the whole-cell extract reaction were also detected after incubation in the nuclear extract; specifically, the splicing reaction displayed a similar time course and dependence on Mg²⁺ and ATP (Fig. 1a). Moreover, as has been reported previously^{15,21}, the

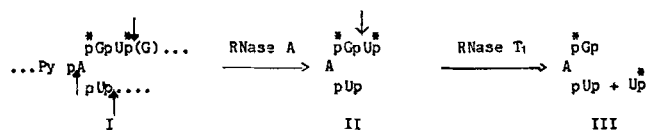
efficiency of RNA processing was significantly higher in the nuclear extract than in the whole-cell extract (data not shown).

Oligonucleotide branch structure

To analyse the base composition of the nuclease-resistant structure in the lariat RNAs, the excised intervening sequence, IVS-A RNA, was isolated from a splicing reaction in which the pre-RNA was labelled with all four [α - 32 P]-labelled nucleoside triphosphates (NTPs) (Fig. 1a). IVS-A RNA was digested to completion with nuclease P₁ and the products separated by two-dimensional TLC (Fig. 1b). In addition to the four anticipated nucleotides, a nuclease-resistant oligonucleotide was resolved, purified and then digested with snake venom phosphodiesterase, yielding radiolabelled pA, pG and pU but no pC (Fig. 1c). Thus, the nuclease P₁-resistant oligonucleotide is composed of A, G and U residues in approximately equal molar quantities. (The variation in radioactivity between the three mononucleotides in Fig. 1c is due to differences in specific activity.) We have reported previously that β -elimination of the P₁ nuclease-resistant oligonucleotide yielded a 2',3',5' ATP (5'pA₃^{2'p}) residue¹⁹, suggesting that the G and U are linked to the 2' and 3' positions at the RNA branch point (see below).

The lariat structure of the intervening sequence is formed by the joining of sequences at the 5' splice site to a site near the 3' end of the molecule¹⁹. Neither the earlier study nor the above data, however, established that the G at the 5' end of the IVS was retained in the formation of the branch. Definitive identification of this G as part of the branch was provided by an analysis of [α - 32 P]GTP labelled RNA. Pre-RNA and IVS-A RNA were digested in parallel with RNase A and analysed by two-

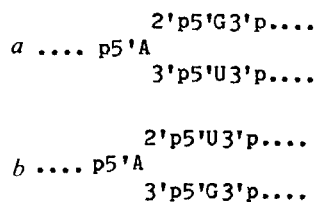
dimensional fingerprint chromatography (Fig. 2*a,b*). A unique branch-containing oligonucleotide was identified in the IVS-A RNA digest that was absent from the pre-RNA digest. Subsequent RNase T₁ digestion of this purified oligonucleotide released a radiolabelled Up and a product migrating in the two-dimensional TLC system with the same mobility as the previously-described RNase T₂-resistant branch trinucleotide (Fig. 2*c* and ref. 19). These results are consistent with the following scheme:



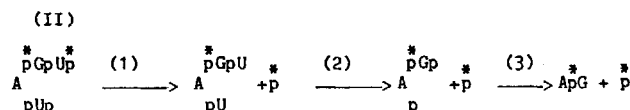
(* Denote ³²P-phosphate residues, ↓ and ↓ indicate cleavage sites of RNases A and T₁, respectively.) The net charge of structures II and III was determined by DEAE TLC; consistent with the above scheme, the RNase A product (II) had a net charge of -7 and treatment with RNase T₁ yielded products with net charges of -6 and -2 (data not shown).

The sequence GUG, established by these data as one arm of the branch, is expected if the 5' splice site is retained precisely in the reaction. As the radioactive phosphate in oligonucleotide III is found in a 2'-5' phosphodiester bond (see below), the phosphate at the 5' splice site must be conserved in branch formation.

Two alternative branch structures are in agreement with the data presented so far, which differ by the distribution of G and U at the 2' or 3' position:



To distinguish between these two possibilities, the RNase A oligonucleotide containing the branch site (structure II) was isolated from [³²P]GTP-labelled IVS-A RNA (Fig. 2*b*) and treated consecutively with phosphatase, periodate/analine (β-elimination), then phosphatase. This sequence of reactions should yield the following set of labelled products:



(* Denote ³²P-phosphate residues.) Structure *a* would yield the dinucleotide A2'p⁵G, whereas structure *b* would yield A3'p⁵G, with a ³²P-phosphate released in both cases. As shown in Fig. 2*d*, the product of the reaction co-migrates in two-dimensional TLC with the marker A2'p⁵G and not with the marker A3'p⁵G, thus the 5' terminal G residue of the intervening sequence is joined via a 2'-5' phosphodiester bond to an A residue.

Location of branch site

The branch site is in the RNase T₁-resistant oligonucleotide UCAUACUUAUCCUG, within the intervening sequence from -19 to -32 nucleotides from the 3' splice site¹⁹. This oligonucleotide contains two AU dinucleotides (underlined in the sequence above) consistent with the established structure of the branch. To discriminate between the two alternative branch sites, we took advantage of the fact that one of the above AU sequences

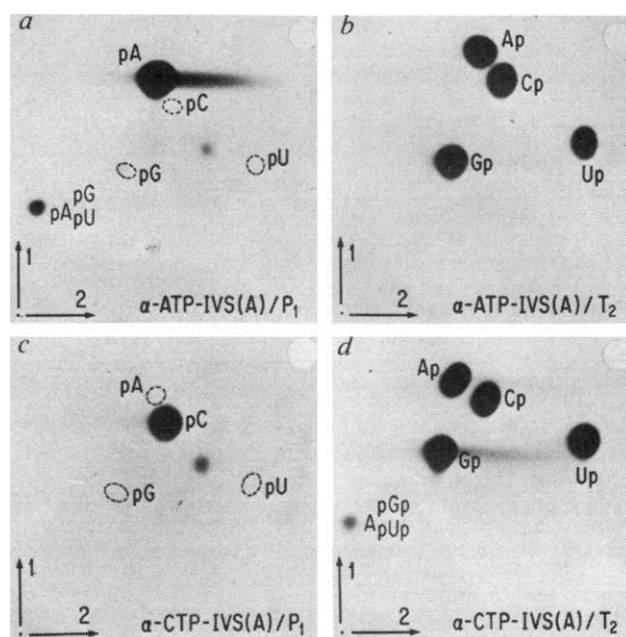


Fig. 3 Localization of the branch site by nearest-neighbour analysis. IVS-A RNAs labelled with either [³²P]-ATP (*a, b*) or [³²P]-CTP (*c, d*) were digested with nuclease P₁ (*a, c*) or RNase T₂ (*b, d*). Products of the reactions were resolved by TLC on cellulose plates in solvent A in the first dimension and solvent B in the second dimension. Positions of nucleoside 5' or 3' monophosphate markers and the branched trinucleotide are indicated. The additional spot observed near the centre of plates *a* and *c* is cytosine 5' phosphate, 2',3'-cyclic phosphate, originating probably from the nonspecific breakdown of RNA (data not shown). Digestions with RNase T₂ (Calbiochem, 1 U in 50 mM ammonium acetate pH 5.3) were for 1 h at 50 °C. Digestions with nuclease P₁ were as in Fig. 1 legend.

is followed by an A whereas the other is followed by a C residue. A phosphate transfer analysis was thus performed by digesting IVS-A RNA labelled with [³²P]-ATP or [³²P]-CTP in parallel with RNase T₂ (Fig. 3). As the branched trinucleotide was labelled with [³²P]-CTP (Fig. 3*d*) but not with [³²P]-ATP (Fig. 3*b*), the downstream AU sequence is the exclusive site of branch formation. Nuclease P₁ digestion was performed as a control for specific labelling of the IVS-A RNAs and to demonstrate the presence of a branch trinucleotide in the [³²P]-ATP-labelled RNA (Fig. 3*a,c*).

As a further demonstration of branch location, the branch containing RNase T₁ oligonucleotide was isolated from IVS-A RNA labelled with both [³²P]-CTP and [³²P]-UTP and digested subsequently with the A-specific RNase U₂. RNase U₂ cleavage requires a 2'-hydroxyl group and therefore will not cleave after the A residue at the branch site. RNase U₂ digestion of the branch containing RNase T₁ oligonucleotide produced UCA, UA and a large branch-containing oligonucleotide CUUA^G_{UCCUG} (Fig. 4*a*). A control digestion of the equivalent RNase T₁ oligonucleotide isolated from pre-RNA gave the expected products UCA, UA, CUUA and UCCUG (Fig. 4*b*). This and the above analyses demonstrate that the branch containing RNase T₁ oligonucleotide has the structure UCAUACUUA^G_{UCCUG}.

Structure of the first exon intermediate

We have shown previously that the IVS-L2-A RNA is a lariat RNA species composed of the intervening sequence and the L2 exon^{18,19}. This RNA appears early in the splicing reaction and is probably part of an intermediate (ref. 18 and Fig. 1*a*). The IVS-L2-A RNA results from a cleavage at the 5' splice site of precursor RNA^{18,19}; the same event should produce the L1 RNA, recoverable as a free 5' exon. An RNA species of the appropriate length (~120 nucleotides) appeared early in the reaction in an

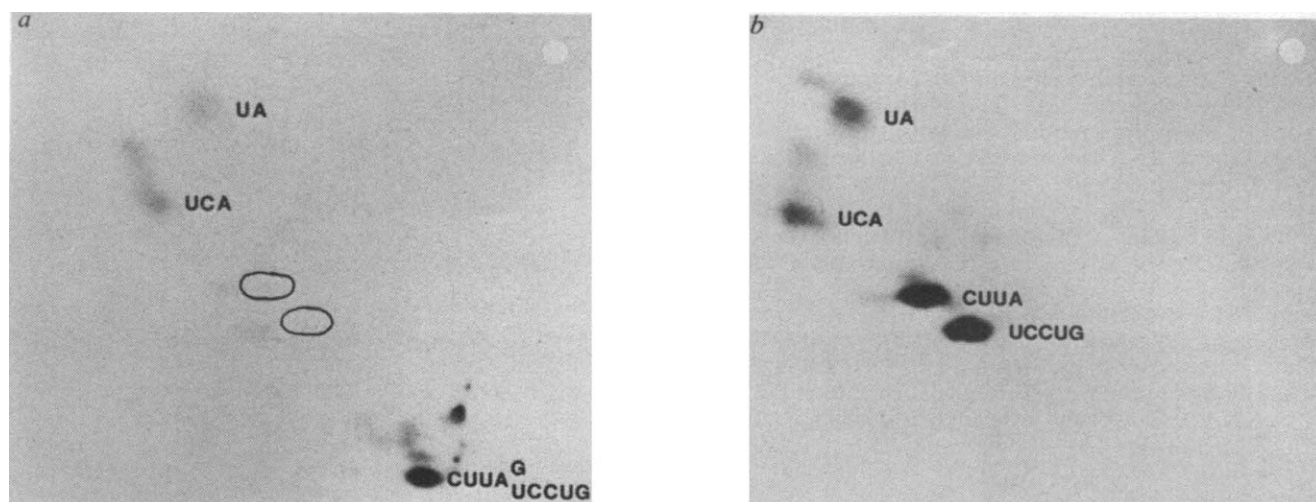


Fig. 4 Localization of the branch site by RNase U_2 analysis. Two-dimensional RNA fingerprint analysis of RNase U_2 digestions of: *a*, the branch containing RNase T_1 oligonucleotide from IVS-A RNA, and *b*, the analogous, unmodified RNase T_1 oligonucleotide from pre-RNA. IVS-A RNA and pre-RNA were labelled with both [α - 32 P]-CTP and [α - 32 P]-UTP. The RNase T_1 oligonucleotides were isolated from a 15% polyacrylamide 8 M urea gel¹⁹ and digested with 0.01 U RNase U_2 (Calbiochem) in 20 mM sodium citrate pH 3.5 for 30 min at 30 °C. Two-dimensional fingerprint chromatography was as in Fig. 2.

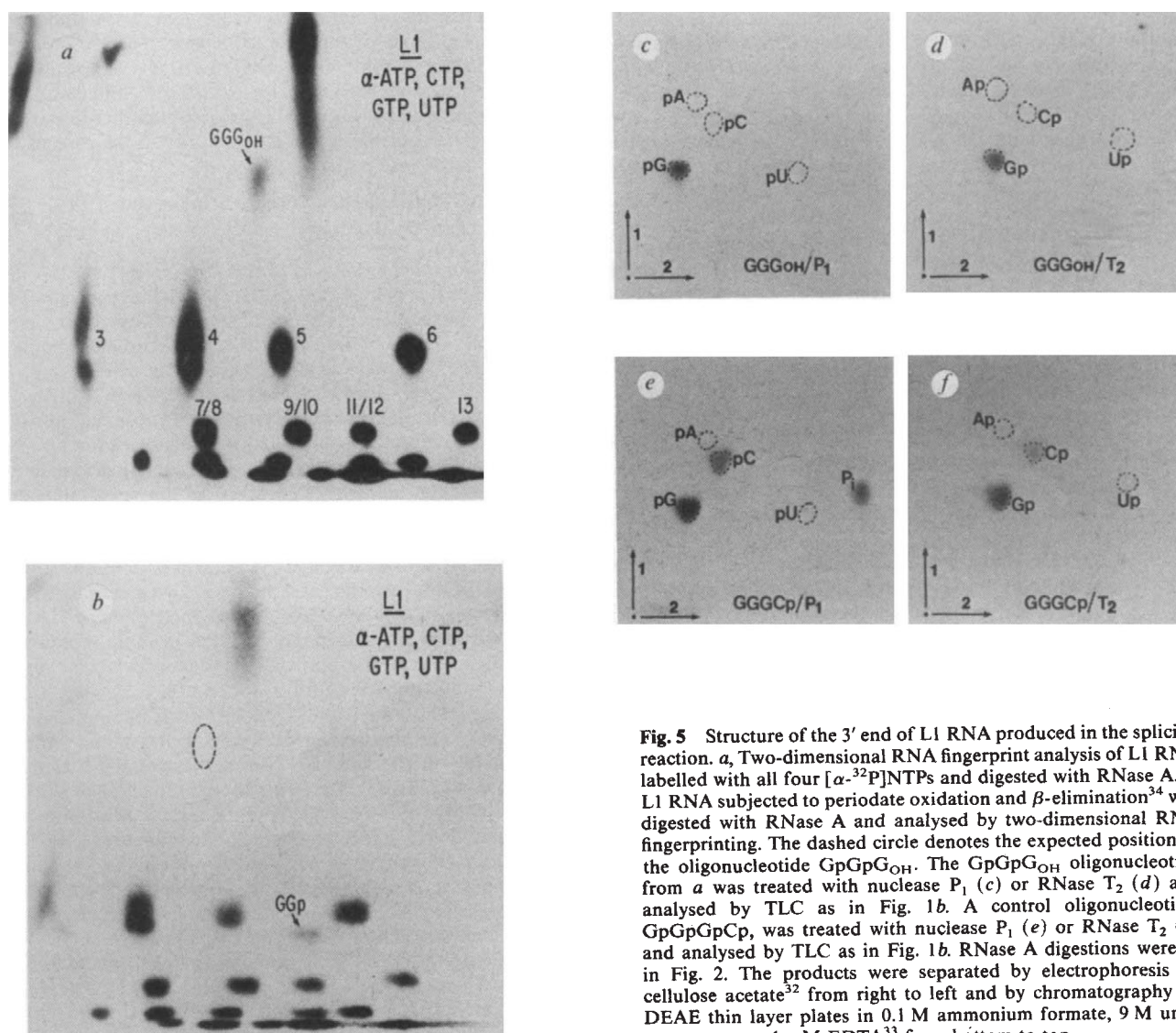


Fig. 5 Structure of the 3' end of L1 RNA produced in the splicing reaction. *a*, Two-dimensional RNA fingerprint analysis of L1 RNA labelled with all four [α - 32 P]NTPs and digested with RNase A. *b*, L1 RNA subjected to periodate oxidation and β -elimination³⁴ was digested with RNase A and analysed by two-dimensional RNA fingerprinting. The dashed circle denotes the expected position of the oligonucleotide GpGpGOH. The GpGpGOH oligonucleotide from *a* was treated with nuclease P_1 (*c*) or RNase T_2 (*d*) and analysed by TLC as in Fig. 1*b*. A control oligonucleotide, GpGpGpCp, was treated with nuclease P_1 (*e*) or RNase T_2 (*f*) and analysed by TLC as in Fig. 1*b*. RNase A digestions were as in Fig. 2. The products were separated by electrophoresis on cellulose acetate³² from right to left and by chromatography on DEAE thin layer plates in 0.1 M ammonium formate, 9 M urea, 1 mM EDTA³³ from bottom to top.

ATP-dependent fashion (Fig. 1a). This RNA contains L1 sequences, shown by RNase A fingerprint analysis and hybridization to single-stranded M13 DNAs containing sequences complementary to the 5' portion of the pre-RNA (data not shown).

As the 3' end of the L1 RNA is involved potentially in the formation of the spliced product, we sought to determine its structure. The L1 RNA cleaved precisely at the 5' splice site should have a 3' sequence of ...UGGG (see Fig. 2e). As shown above, the phosphate moiety from the 5' splice site is contained in the 2'-5' phosphodiester bond of the branch. Thus, it is probable that the L1 RNA has a 3' hydroxyl terminus. To investigate this, an RNase A digestion of the purified L1 RNA was analysed by a two-dimensional fingerprint. A unique oligonucleotide with the anticipated mobility of GpGpG_{OH} is present in L1 RNA labelled with four [α -³²P]NTPs (Fig. 5a) or [α -³²P]GTP (data not shown). This oligonucleotide is not present in L1 RNA labelled with [α -³²P]CTP. It is also absent from both L1-2 RNA and pre-RNA labelled with all four [α -³²P]NTPs (data not shown). If the L1 RNA terminates with 2' and 3' hydroxyl groups, then periodate oxidation and β -elimination should yield GpGp as a product of RNase A digestion. As anticipated, the GpGpG_{OH} oligonucleotide (net charge -2) was absent and a new oligonucleotide migrating at the expected position of GpGp (net charge -3) appeared after this treatment (Fig. 5b), consistent with the location of GpGpG_{OH} at the 3' terminus of L1 RNA. To confirm the identity of the GpGpG_{OH} oligonucleotide, we determined its base composition by elution from the fingerprint and digestion with RNase T₂ or nuclease P₁. Only the anticipated products Gp and pG were observed (Fig. 5c,d); control digestions on oligonucleotide GpGpGpCp yielded the expected products (Fig. 5e,f). Note that nuclease P₁ digestion did not release phosphate from oligonucleotide GpGpG_{OH}, consistent with the presence of a 3' terminal hydroxyl group. The assignment of the 3' hydroxyl group in the L1 exon agrees with the previous finding that the phosphate moiety at the splice junction of the two exons is derived from the 3' splice site¹⁹.

Mechanism of mRNA precursor splicing

On the basis of the work reported here and other recent findings^{18,19,21}, we propose a scheme for the splicing of mRNA precursors (Fig. 6). A multicomponent precursor mRNA-ribonucleoprotein complex with a poorly defined structure is formed first; it probably contains a small nuclear U1 ribonucleoprotein particle bound to the 5' splice site^{16,23,24}. An involvement of other small nuclear ribonucleoproteins in mRNA splicing is also possible^{25,26}. The lag in the time course observed *in vitro*¹³⁻¹⁵ and the recognition of the 5' cap structure at an early stage of the reaction¹⁷ may reflect an ATP-dependent organization of precursor into a complex which properly positions the two exons.

The first characterized step in the reaction is cleavage of the RNA precursor at the 5' splice site, producing a 5' exon terminating in a 3' hydroxyl group and a lariat structure containing the remainder of the RNA. The branch in the lariat results from the joining of the 5' terminal G of the intervening sequence to the 2' hydroxyl group of an A. The branch site in this particular example is positioned 24 nucleotides upstream of the 3' splice site. Interestingly, there is a distinct complementarity between the sequences at the 5' terminus of the intervening sequence and the nucleotides preceding the A at the branch site (see Fig. 6). The close proximity of the 2' hydroxyl of the A at the branch site to the 5' splice site in a secondary structure could facilitate the cleavage/ligation reaction. It is not clear whether cleavage at the 5' splice site and formation of the 2'-5' phosphodiester bond occurs through a concerted transesterification reaction⁸ or through a series of enzymatic cleavage/ligation reactions. The phosphate group from the 5' splice site (circled in Fig. 6) is retained in the 2'-5' phosphodiester bond of the branch structure, consistent with the released 5' exon terminating in a 3' hydroxyl group. A second cleavage/ligation reaction joins the two exons by a 3'-5' phosphodiester bond and releases the lariat form of

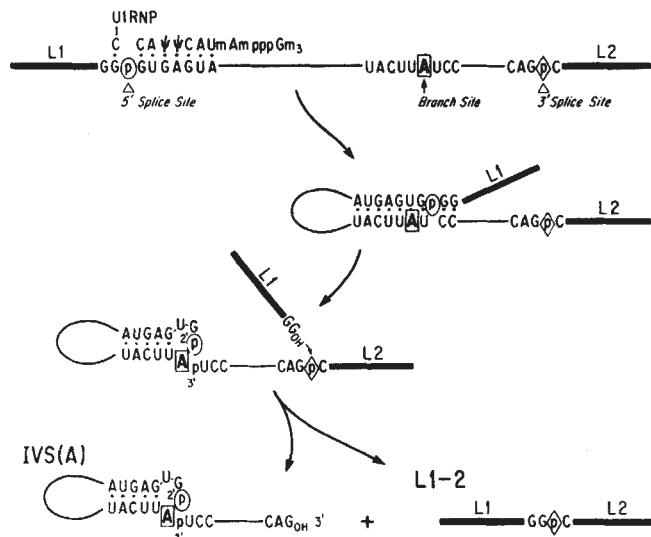


Fig. 6 Proposed scheme of pre-mRNA splicing.

the intervening sequence with a 3' hydroxyl group¹⁹. In this reaction, the phosphate group at the 3' splice site (enclosed by a diamond in Fig. 6) is joined to the 3' hydroxyl group of the 5' exon. As in the previous step, this reaction could either proceed as one concerted transesterification reaction, or could consist of a series of separate reactions.

The degree of complementarity (10 base pairs (bp) including G:U pairing) between sequences at the 5' splice site and the branch site of IVS-1 is probably greater than that found in other introns. The equivalent complementarity in the case of human β -globin is 9 base pairs²¹.

5'ACUAUGGUUG:GA5' 5' splice site of IVS-1

5'AGGCACUGAC UC3' branch site of IVS-1

Some complementarity is found also in yeast mRNA introns, where the conserved sequence UACUAAC is required upstream of all 3' splice sites; the sequence at the 5' splice site is almost invariably G:GUAUGU^{27,28}, these two sequences being complementary²⁸. Single base changes in the yeast intron UACUAAC sequence reduce the efficiency of splicing²⁹; the UACUACC variant is totally defective in splicing, not surprisingly as the mutated A is probably the base of the branch site.

Sequences encompassing potential branch sites in mammalian introns are not as highly conserved as the yeast equivalent. A search of mammalian intron sequences using the yeast paradigm yielded a consensus sequence of CTGAC²⁶. In the rabbit β -globin large intron, deletion of sequences containing the best example of this consensus sequence, almost certainly the normal branch site, did not abolish accurate splicing *in vivo*³⁰. Although alternative branch sites might be used in these mutants, these data suggest that the branch site sequence is not the major determinant for selection of 3' splice sites.

The sequences near the 5' splice site are complementary to both the 5' terminal sequences of U1 RNA and branch-site sequences (see Fig. 6). Base-pairing between the 5' terminus of U1 RNA and the 5' splice site is thought to mediate recognition for splicing; following selection of the 5' splice site, the U1 RNA could be displaced by branch-site sequences. As described above, the efficiencies of splicing may not be highly dependent on the degree of complementarity between the 5' splice site and the branch site. Accurate splicing *in vivo* has been often observed with chimaeric genes, where the 5' splice site and the 3' splice site (including the branch site) originated from two different and unrelated genes (see, for example, ref. 31). The degree of complementarity between the branch site and the 5' splice site may regulate the rate of splicing by stabilizing their interaction. For example, it might determine the order of excision of multiple

introns in a single transcript. In cases where alternative splice sites are used, the relative strengths of the interactions may determine the ratios of splice site use. Until the branch sites used in such RNAs are determined, these possibilities are difficult to evaluate.

The formation of a branch seems to be a unique feature of the splicing of pre-mRNAs. This unconventional structure might be simply a result of the splicing mechanism. It is also possible, however, that the reactive 5' end of the intervening sequence is prevented in this way from recombining with the 5' exon, driving the reaction in the forward direction⁹. Finally, the branch structure might serve as a positive signal for rapid degradation of intervening sequences *in vivo*.

The pathway of splicing described here may not be restricted necessarily to nuclear protein-coding genes. A recent finding that the excised, circular intervening sequence produced during splicing of the transcripts of the yeast mitochondrial gene for subunit I of cytochrome *c* oxidase may be a lariat RNA contain-

ing a branch structure (L. A. Grivell, personal communication) suggests that the splicing mechanism operating in these two classes of genes is similar. Moreover, the fact that nuclear introns contain internal sequences important to splicing (branch sites) relates this class of introns to both mitochondrial mRNA introns and nuclear and mitochondrial rRNA introns (for review see ref. 3). Further studies are necessary to determine whether the structural similarities of these introns are accompanied by a corresponding resemblance in the mechanisms of splicing.

We thank Claire Moore for discussions and for preparation of the nuclear extract. P.J.G. acknowledges Peter Gegenheimer for suggestions on RNA fingerprint analysis. M.M.K. holds a postdoctoral fellowship from the Jane Coffin Childs Memorial Fund, P.J.G. from Helen Hay Whitney Foundation and R.A.P. from the Myron Bantrell Fellowship. This work was supported by NIH grants RO1-GM32467 and PO1-CA26717 and NSF grant PCM-8200309 to P.A.S. and partially by NIH core grant PO1-CA14051.

Received 18 October; accepted 6 December 1984.

- Abelson, J. A. *Rev. Biochem.* **48**, 1035-1063 (1979).
- Sharp, P. A. *Cell* **23**, 643-646 (1981).
- Cech, T. *Cell* **34**, 713-716 (1983).
- Konarska, M., Filipowicz, W. & Gross, H. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1474-1478 (1982).
- Greer, C. L., Peebles, C. L., Gegenheimer, P. & Abelson, J. *Cell* **32**, 537-546 (1983).
- Filipowicz, W. & Shatkin, A. J. *Cell* **32**, 547-557.
- Laski, F. A., Firc, A. Z., RajBhandary, U. L. & Sharp, P. A. *J. biol. Chem.* **258**, 11974-11980 (1983).
- Cech, T. R., Zaug, A. J. & Grabowski, P. J. *Cell* **27**, 487-496 (1981).
- Kruger, K. *et al. Cell* **31**, 147-157 (1982).
- Zaug, A. J., Grabowski, P. J. & Cech, T. R. *Nature* **301**, 578-583 (1983).
- Kole, R. & Weissman, S. M. *Nucleic Acids Res.* **10**, 5429-5445 (1982).
- Padgett, R. A., Hardy, S. F. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5230-5234 (1983).
- Hernandez, N. & Keller, W. *Cell* **35**, 89-99 (1983).
- Hardy, S. F., Grabowski, P. J., Padgett, R. A. & Sharp, P. A. *Nature* **308**, 375-377 (1984).
- Krainer, A. R., Maniatis, T., Ruskin, B. & Green, M. R. *Cell* **36**, 993-1005 (1984).
- Padgett, R. A., Mount, S. M., Steitz, J. A. & Sharp, P. A. *Cell* **35**, 101-107 (1983).
- Konarska, M. M., Padgett, R. A. & Sharp, P. A. *Cell* **38**, 731-736 (1984).
- Grabowski, P. J., Padgett, R. A. & Sharp, P. A. *Cell* **37**, 415-427 (1984).
- Padgett, R. A., Konarska, M. M., Grabowski, P. J., Hardy, S. F. & Sharp, P. A. *Science* **225**, 898-903 (1984).
- Wallace, J. C. & Edmonds, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 950-954 (1983).
- Ruskin, B., Krainer, A. R., Maniatis, T. & Green, M. R. *Cell* **38**, 317-331 (1984).
- Dignam, J. D., Lebowitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
- Mount, S. M., Pettersson, I., Hinterberger, M., Kärner, A. & Steitz, J. A. *Cell* **33**, 509-518 (1983).
- Krämer, A., Keller, W., Appel, B. & Lührmann, R. *Cell* **38**, 299-307 (1984).
- Ohshima, Y., Itoh, M., Okada, N. & Miyata, T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4471-4474 (1981).
- Keller, E. B. & Noon, W. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7417-7420 (1984).
- Langford, C. J. & Gallwitz, D. *Cell* **33**, 519-527 (1983).
- Pikielny, C. W., Teem, J. L. & Rosbash, M. *Cell* **34**, 395-403 (1983).
- Langford, C. J., Klinz, F.-J., Donath, C. & Gallwitz, D. *Cell* **36**, 645-653 (1984).
- Wieringa, B., Hofer, E. & Weissmann, C. *Cell* **37**, 915-925 (1984).
- Chu, G. & Sharp, P. A. *Nature* **289**, 378-382 (1981).
- Silberklang, M., Gillum, A. M. & RajBhandary, U. L. *Meth. Enzym.* **59**, 58-109 (1979).
- Konarska, M., Filipowicz, W., Domdey, H. & Gross, H. J. *Nature* **293**, 112-116 (1981).
- Rose, J. K. & Lodish, H. F. *Nature* **262**, 32-37 (1976).

LETTERS TO NATURE

Pavo XD-10, an X-ray QSO with extended optical structure

R. Gilmozzi[†]*, J. V. Wall[†], P. G. Murdin[†],
P. R. Jorden[†], D. J. Thorne[†], I. G. van Breda[†]
& J. A. Peacock[‡]

* ESA Satellite Tracking Station, Villafranca del Castillo,
PO Box 54065, Madrid, Spain

† Royal Greenwich Observatory, Herstmonceux Castle, Hailsham,
East Sussex BN27 1RP, UK

‡ Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

Deep charge-coupled device (CCD) images of the quasi-stellar object (QSO) identified with the Einstein X-ray source Pavo XD-10 show an optical structure that is dramatically extended. Here we discuss the morphology, the photometry from the CCD data, and the limited spectral information which indicates a redshift (z) of 0.72 for the system. We show that if the extended structure represents a galaxy hosting the $z = 0.72$ QSO, the galaxy luminosity is 2 mag brighter than the host galaxies detected for QSOs at lower redshifts. Instead, the system may be the result of a galaxy-galaxy interaction. A third and more speculative possibility is that it is the result of gravitational lensing, with much of the 'underlying' structure representing the lens galaxy in the line-of-sight.

Pavo XD-10 was identified as an X-ray QSO in a detailed investigation¹ of the Pavo field in the Einstein Observatory deep X-ray survey². To extend this investigation, we took broad-band (BVRI) exposures of several areas in the Pavo field at the prime focus of the 3.9-m Anglo-Australian Telescope (AAT) in September 1981, using a CCD camera built at the Royal Greenwich Observatory (RGO)³. We noted then the extension to QSO

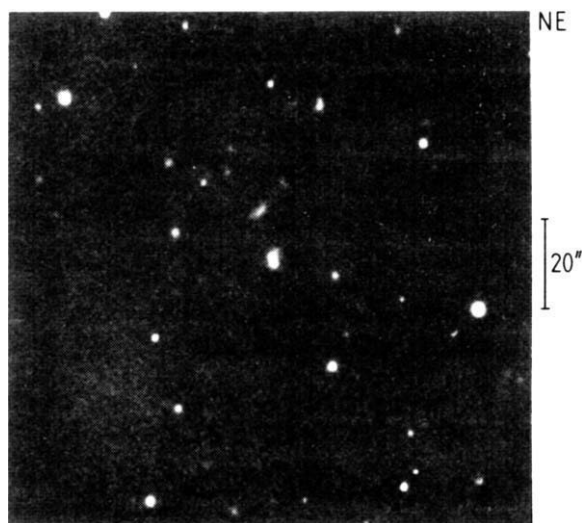


Fig. 1 A photographic representation of the region about Pavo XD-10 (1950.0 $\alpha = 21$ h 11. min 32.6 s, $\delta = -67^\circ 54' 26''$), obtained from the 3,600-s I-band image taken with a CCD camera at the Cassegrain focus of the Danish 1.5-m telescope at ESO/La Silla. The QSO identification is the image elongated north-south in the centre of the frame, object 'a' of ref. 1.

XD-10. Much better observing conditions (photometric, with 1 arc s seeing) were obtained during the commissioning of a second RGO CCD camera at the $f/8$ Cassegrain focus of the Danish 1.5-m telescope at the European Southern Observatory (ESO)/La Silla in August 1982. Both CCD cameras were based on RCA 53612 detectors cooled to 150 K in N_2 cryostats. The

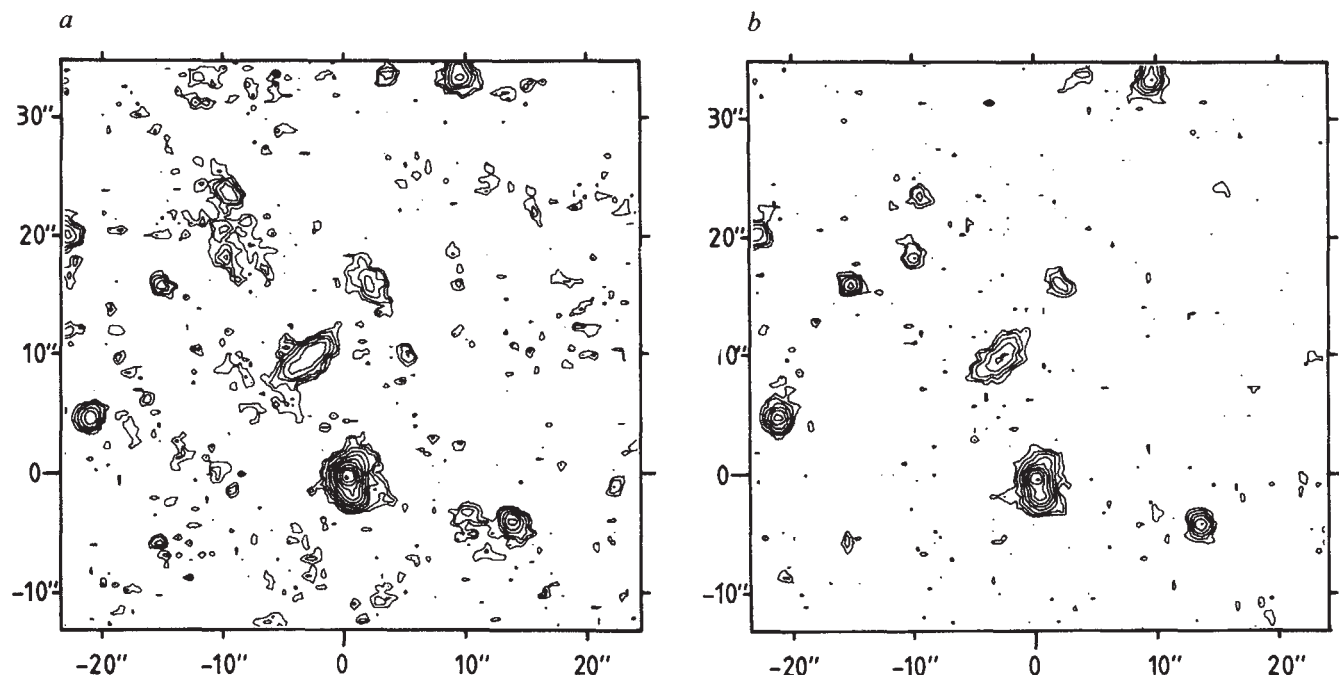


Fig. 2 Contour plots from the cleaned CCD images of the 50×50 arc s field about Pavo XF-10; *a*, *R*-band, 3,600-s; *b*, *I*-band, 3,600-s Danish 1.5-m telescope. The cleaning process included removal of DC offset, flat-fielding, patching of a few cosmic-ray events and smoothing with a gaussian filter of $\sigma = 0.5$ pixels. The axes indicate arc s in RA and dec. with origins at the optical image identified with Pavo XD-10 (image 'a' of ref. 1). The contours are at intervals of $0.5 \text{ mag arc s}^{-2}$, starting at $23 \text{ mag arc s}^{-2}$ in *I* and $25 \text{ mag arc s}^{-2}$ in *R*, these representing about 1% of sky brightness in each passband.

images (512×320 pixels, each 0.49×0.49 arc s at the AAT prime focus, 0.47×0.47 arc s at the 1.5-m Cassegrain focus) were obtained in a *BVRI* system calibrated by observations of a sequence near NGC288⁴ and near NGC300⁵. With the AAT we used exposure times of 240 s in *B* and *V*, and 400 s in *R* and *I*, whereas with the 1.5-m telescope, we took exposures of 3,600 s each in *R* and *I*. Analysis was carried out at the RGO node of the STARLINK computer network (SERC).

The single optical spectrum available¹ is of low signal/noise ratio, showing a blue continuum and a single emission line at $4,800 \text{ \AA}$. If this line is $\text{Mg II } \lambda 2798$, then $z = 0.72$; a second, but less probable, possibility is $\text{C III] } \lambda 1909$, $z = 1.51$. A possible broad absorption feature is at $5,180 \text{ \AA}$.

A photographic representation of the region of the *I*-frame is shown in Fig. 1, whereas Fig. 2 shows contour maps in *R* and *I*. The QSO appears embedded in a nebulosity, which is extended to the south-east and in a curve to the north-east. This structure suggests neither spiral nor elliptical galaxy morphology. In the *I* observations, the extended structure to the south peaks as a second maximum (Fig. 2*b*), the separation of the peaks being 1.6 arc s.

Table 1 shows the results of our photometry. The magnitude of the total system was determined from simple aperture photometry. In the *R* and *I* frames, the nucleus of the system, designated as 'QSO' in Table 1, has the same profile as that of the point-spread function determined from stars in these frames. Accordingly, we determined the magnitude of the nuclear component from the height of this (stellar) profile above the nebulos-

ity; the magnitude of the nebulosity was then estimated from the total intensity less that of the stellar nucleus, the 'QSO'. The total *R* and *I* magnitudes, measured from 1.5-m telescope data obtained in photometric conditions, have r.m.s. errors of 0.1 mag. The *B* and *V* magnitudes from the AAT observations have larger errors of 0.2 mag, some of which is due to signal/noise and to errors in background estimation as for the *R* and *I* magnitudes, and some to uncertainty in zero-point determination. Conditions were not photometric, and to set zero points we plotted $(B - V)$ against $(V - R)$ for the 12 brightest stars in the frame. These are predominantly G stars and form a well-defined locus in the two-colour plane, from which, given the *R* magnitudes from the 1.5-m data, zero-points in *B* and *V* could be estimated. These limited *BVRI* colour data (Table 1) demonstrate that the nebulosity is redder than the stellar core, which dominates the light of the system in *B* and *V*.

The system may be associated with a faint group of galaxies whose colours and magnitudes are consistent with ellipticals at $z \approx 0.7$. Figure 1 shows several extended images just north-west of Pavo XD-10, including a disk galaxy 11 arc s north-west. Using aperture photometry on the CCD images, magnitudes in *R* and *I* were measured for nine of these objects. From these (and excluding the disk galaxy), $\bar{R} = 21.1 \pm 0.1 \text{ mag}$ and $\bar{R} - \bar{I} = 1.2 \pm 0.1 \text{ mag}$, where the uncertainties are r.m.s./ $\sqrt{N-1}$. The assumption of a typical elliptical luminosity of $B^* = -20.6 \text{ mag}$ ($H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$)⁶ and $(R - I)$ colours for elliptical galaxies as a function of redshift, computed by C. R. Benn and G. Grueff (personal communication) from established spectral-energy distributions^{7,8}, yields an estimate for the distance corresponding to $z = 0.7$. The mean colour is consistent with ellipticals at $0.3 < z < 0.8$. The group is ~ 60 arc s in extent, corresponding to $\sim 500 \text{ kpc}$ at $z = 0.7$.

These results permit several interpretations. First, the QSO is in a galaxy, which itself is located in a group or cluster. Following the approach of Wyckoff *et al.*⁹, the angular extent of the radial profile at $R = 26 \text{ mag arc s}^{-2}$ is $\theta = 8 \text{ arc s}$. At $z = 0.72$, the system lies on the extrapolated plot of $\log cz$ against $\log \theta$ obtained by Wyckoff *et al.* in their survey of structure in 15 low-redshift QSOs. If this interpretation is correct, and $z = 0.72$ is confirmed,

Table 1 Magnitudes for Pavo XD-10

Magnitudes in:	<i>B</i>	<i>V</i>	<i>R</i>	<i>I</i>
QSO + nebulosity	19.6	19.5	18.9	18.0
QSO			19.6	19.1
Nebulosity			19.7	18.5

Seeing conditions in *B* and *V* frames did not permit separation of stellar and nebulous components.

Pavo XD-10 represents the most distant QSO for which associated nebulosity is well resolved (although Gehren *et al.*¹⁰ report some underlying structure for a QSO at $z = 0.83$). However, the nebulosity lies well below the Wyckoff *et al.* plot of $\log \theta$ against $\log R$, and well above their line for the brightest cluster galaxies in the $\log cz$ against R_{26} (corrected) plot. The host galaxy appears to be too bright by at least 1 mag.

Suppose instead, then, that the nebulosity represents a galaxy-galaxy interaction, as suggested by its bent shape and the presence of two nuclei in the I image (Fig. 2b). Others^{11,12} have drawn attention to the importance of interactions in the QSO phenomenon. If two galaxies are involved, the total angular extent and magnitude are not representative of a single object, and it is then plausible for the system (two galaxies, the nucleus of one having flared to become a QSO) to have a redshift as high as $z = 0.72$.

A third possibility is that the QSO and the nebulosity are completely independent, the QSO at $z = 0.72$, the nebulosity a galaxy near the line-of-sight at a redshift of 0.2–0.4. The colour ($R - I = 1.2$) and magnitude ($R = 19.7$) are consistent with an elliptical or early spiral (Sab) of moderate luminosity at such a distance; the weak absorption feature in the spectrum (of very marginal significance) at 5,180 Å could represent the Ca II H and K break at $\lambda 4,000$, $z = 0.30$. Considerable gravitational amplification of the QSO image is probable with such a configuration. There are two simple possibilities: the two peaks in the

I image either represent two principal images or represent one principal image and the galaxy nucleus, with other images perhaps obscured by the central region of the galaxy. On the first model, assuming that the galaxy is an isothermal sphere¹³, the observed separation of 1.6 arc s implies a line-of-sight velocity dispersion σ_v of 240 km s⁻¹ and an intrinsic separation (QSO-galaxy nucleus) of ≤ 0.8 arc s; on the second model, the intrinsic separation in arc s is $1.6 - 0.8(\sigma_v/240)^2$. For the first model, the QSO would be amplified by a factor > 2 . We can only avoid significant amplification if the galaxy has little mass ($\sigma_v < 150$ km s⁻¹), in which case the second model of only one image would apply. Note, however, that the gravitational potential of any cluster around the galaxy aids the lensing process (as with the double QSO 0957 + 56). Finally, if the QSO is small enough ($\sim 10^{-3}$ pc), individual stars in the galaxy can act as lenses. With light from the QSO passing 1.6 arc s from the centre of an isothermal galaxy, the probability of large amplification (> 2) is $\sim (\sigma_v/600)^2$ (ref. 13), which is again appreciable for a moderately massive galaxy. If we are indeed seeing an intervening galaxy, the small galaxy-QSO separation makes strong lensing difficult to avoid.

We thank the Danish time allocation committee for use of the 1.5-m telescope, Henning Jorgensen for generous assistance during the observations and Sue Worswick and Richard Ellis for very helpful discussions. R.G. is on leave from Istituto Astrofisica Spaziale, CNR, CP67, 0044 Frascati, Italy.

Received 26 September; accepted 19 December 1984.

1. Griffiths, R. E. *et al. Astrophys. J.* **269**, 375–386 (1983).
2. Giacconi, R. *et al. Astrophys. J. Lett.* **234**, L1–L7 (1979).
3. Jorden, P. R., Thorne, D. J. & van Breda, I. G. *Proc. Soc. photo-opt. instrum. Engr.* **331**, 87–95 (1982).
4. Hawkins, M. R. S. *Mon. Not. R. astr. Soc.* **194**, 1013–1023 (1981).
5. Graham, J. A. *Publ. astr. Soc. Pacif.* **93**, 29–31 (1981).

6. Ellis, R. S. in *The Origin and Evolution of Galaxies* (eds Jones, B. J. T. & Jones, J. E.) 255–293 (Reidel, Dordrecht, 1983).
7. Pence, W. D. *Astrophys. J.* **203**, 39–51 (1976).
8. Coleman, G. D., Wu, C. C. & Weedman, D. W. *Astrophys. J. Suppl.* **43**, 393–416 (1980).
9. Wyckoff, S., Wehinger, P. A. & Gehren, T. *Astrophys. J.* **247**, 750–761 (1981).
10. Gehren, T., Fried, J., Wehinger, P. A. & Wyckoff, S. *Astrophys. J.* **278**, 11–27 (1984).
11. Hutchings, J. B. & Campbell, B. *Nature* **303**, 584–588 (1983).
12. Hutchings, J. B., Crampton, D. & Campbell, B. *Astrophys. J.* **280**, 41–50 (1984).
13. Peacock, J. A. *Proc. 24th Liège Symp.*, 86–92 (Institut d'Astrophysique, Université de Liège, 1983).

Low-frequency fluctuations of the electric field in the equatorial ionosphere

V. L. Patel* & Pablo Lagos†

* Department of Physics, University of Denver, Denver, Colorado 80208, USA

† Radio Observatorio de Jicamarca, Instituto Geofísico del Perú, Apartado 3747, Lima 100, Perú

High time-resolution radar measurements of ionospheric electric fields have recently permitted studies of low-frequency hydromagnetic waves, but most of the research has depended on electron drift velocity from data obtained from the Scandinavian Twin Auroral Radar Experiment (STARE), taken at high latitudes ($\lambda_m \sim +70^\circ$) and used to study oscillations with periods of 200–300 s. At high latitudes, the low-frequency pulsations provide important information on the magnetospheric plasma process^{1–3}, but, at low latitudes, coupling of the ionosphere and magnetosphere remain unexplored. The purpose of our study is to examine the fluctuations in the electric field on a time scale longer than ion cyclotron periods, $T_{ci} \sim 0.6$ s (ref. 4), or for frequencies $\omega \ll \omega_{ci} \sim 10^2$ rad s⁻¹ (the ion cyclotron frequency in the equatorial ionosphere), which have not been studied at low latitudes. We report here observations of these waves with periods larger than 45 s from the high time-resolution data taken at Jicamarca, Perú.

The Jicamarca Radar Observatory near the geomagnetic equator (magnetic dip $\sim 2^\circ$) at 75° W longitude provides observations of ionospheric electron drift velocities through the incoherent scatter technique⁵. The electric field is determined from the electron drift velocity, $V_e = E \times B/B^2$ (where E and B are the electric and magnetic field vectors, respectively), which

in the equatorial ionosphere is approximately vertical under the influence of the east-west electric field. The average characteristic of the equatorial ionospheric electric field and its long-duration variations have been reviewed by Balsley⁶.

The observations of electron drift velocity were made on 10 October 1981 beginning at 14:21 local time (75° W). The data were obtained at the initial altitude of 200 km and then at height increments of 15 km, of which there were 50, so that the highest altitude was 935 km. We have analysed data from the antenna A, 3.4° to the west. To filter high-frequency noise or fluctuations, the original 3.0-s-interval data were smoothed by a 10-point smoothing/averaging process and the results over 20 min duration are shown in Fig. 1 for altitudes of 200–935 km. The quasi-periodic fluctuations with 90–120-s periods ($\omega \sim 0.01$) are evident throughout, but particularly in the 4–8-min and 10–16-min intervals. These fluctuations approximately represent the vertical component of an oscillation of the electrons in the ionosphere (estimated amplitude varying from 250 to 500 m) under the influence of the east-west electric field component. Similar fluctuations are apparent in the data processed for a further 100 min. To obtain quantitative information, we have performed a power spectrum analysis of the drift velocity data, supplemented by an additional 20-min data set (14:41–15:01 local time). The spectra were obtained from the 3.0-s data for the 10 heights shown in Fig. 1. The result for one altitude of 245 km is presented in Fig. 2. The predominant frequencies are 0.01, 0.008 and 0.014 Hz. As the data segment is 20 min, we do not consider the very low-frequency peak at 0.002 Hz reliable, but a frequency in the range 0.002–0.004 Hz is always present. At all 10 altitudes, the highest power occurs between 0.008 and 0.0126 Hz, that is, wave activity is present of 80–125-s periodicity, and, in this frequency range, we do not find any systematic change in power with altitude. The visual and power spectrum

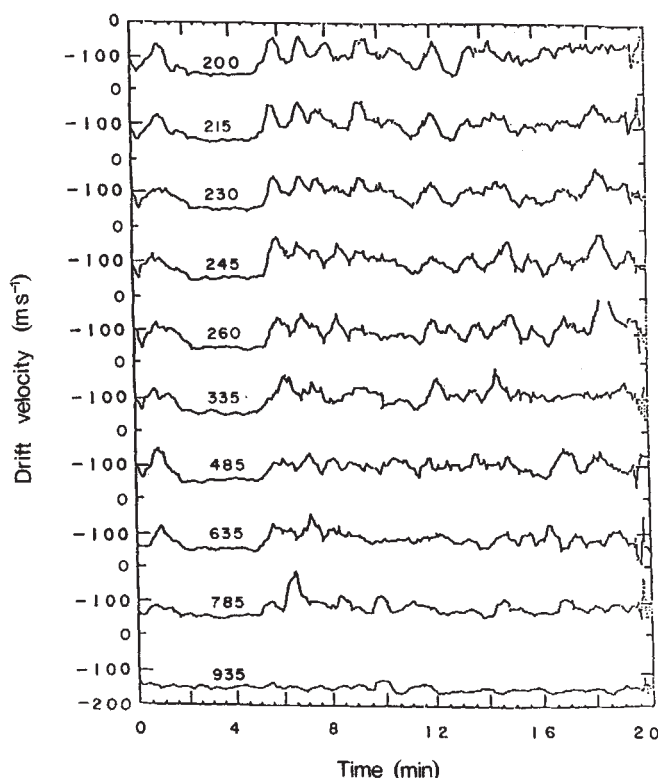


Fig. 1 Electron drift velocity versus time. The zero time is 15:01 local time at 75° W. The data are from antenna A. The number on each plot designates the altitude in km. The scale for the drift velocity is 100 m s⁻¹ for each division, the tick marks beginning with zero and -100 etc. for each plot.

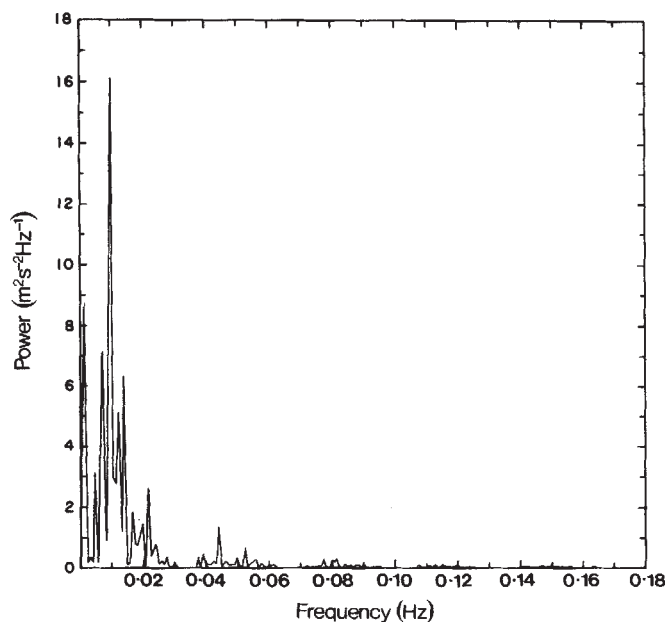


Fig. 2 Power spectrum of electron drift velocity at an altitude of 245 km. The data used in the analysis are from antenna A during 14:41–15:01 local time.

analysis of the electron drift velocity data establish the existence of low-frequency ($\omega < \omega_{ci}$) waves in the electric field.

Although the analysis of the indirect observations of electric field alone does not allow a unique determination of the nature or source of the oscillations, we may identify various possible interpretations. The low-frequency fluctuations in the electric field may be associated with electrostatic or hydromagnetic modes. First, we consider the electrostatic mode. In a plasma with density or temperature gradients, low-frequency drift waves can be generated⁷. At the low latitudes and ionospheric altitudes relevant to our observations, the electron density and temperature show variations that indicate inhomogeneities in these plasma parameters^{8,9}. The ionospheric plasma is a low- β plasma, where β is a ratio of plasma pressure to magnetic pressure. For example, $\beta = 8\pi nkT/B^2 \sim 10^{-4}$, if we take the plasma density $n = 10^6 \text{ cm}^{-3}$ at 300 km over the geomagnetic equator, the plasma temperature $T = 2,000 \text{ K}$ and $B = 0.244 \text{ G}$, extrapolated from a surface field of 0.28 G (k is Boltzmann's constant). The plasma parameters are taken from Brace *et al.*⁸ with the assumption that $n_e \approx n_i = n$, $T_e = T_i = T$, where e , i refers to electrons and ions, respectively. In a low- β plasma, the drift waves are primarily electrostatic (see ref. 7 for the orientation of field perturbations). If we assume that the density gradient in the vertical (radial) direction and the wave vector, k , in the azimuthal direction are approximately perpendicular to B , the electric field perturbations will be parallel to k in the azimuthal direction. The component of the observed electric field fluctuations may be the drift mode driven by internal energy sources from the plasma gradients.

Second, we consider the interpretation of these fluctuations as hydromagnetic modes, which have been studied by Tanenbaum and Mintzer¹⁰. If we consider a simple model of a uniform magnetic field in a uniform plasma, the transverse and longitudinal mode are uncoupled. For the transverse mode, $k \perp B$, the wave electric perturbation will be in a plane perpendicular to the meridian plane and perpendicular to B , that is, in a plane

defined by azimuthal and radial directions. If we assume standing waves along the magnetic field, $\omega = k_{\parallel} V_A$, where $V_A = B/\sqrt{4\pi\rho}$ in c.g.s. units. We calculate the plasma density at the ionospheric altitude of 300 km at low latitudes and compare it with the observations. For waves with period of 90 s and $k_{\parallel} \approx 1/R$, where $R = 6,471 + 300 = 6,771 \text{ km}$, we obtain V_A and then ρ from the value of B cited above. Assuming a pure hydrogen plasma with $\rho = nm_p$, where m_p is the proton mass, we obtain $n \approx 10^6 \text{ cm}^{-3}$, but if we include heavy ions, the plasma density will be lower than this estimate. The estimate of n from this simple model is not inconsistent with the observations of Brace *et al.*⁸.

For the longitudinal mode, $k \perp B$ and k in the vertical (radial) direction, the wave perturbation in the electric field will be in a plane perpendicular to the meridian (B, k) plane with $k \cdot E_{\text{wave}} = 0$ and may appear in a component of east-west electric field. In this mode, it is possible that waves have been generated outside the ionosphere, that is, in the magnetosphere. The energy source in the magnetosphere requires propagation of energy perpendicular to the magnetic field, that is, in the radial direction from the magnetosphere, so as to reach the low-latitude ionosphere. Recently, the Poynting flux in the radially-inward direction has been observed by the geosynchronous satellite GEOS 2 in low-frequency mode¹¹ indicating that the low-frequency oscillations in the ionospheric electric field may have their origin in the magnetosphere. Further studies of these fluctuations may provide important information on the problem of low-latitude ionospheric-magnetospheric coupling. The low-frequency waves, when coupled to other modes, can provide a mechanism for energy transfer in the ionosphere⁴.

We thank Richard Ludlow and Vic Patel for assistance in data reduction. This research has been supported by NSF grant ATM-84-07284.

Received 6 September; accepted 5 December 1984.

1. Walker, A. D. M., Greenwald, R. A., Stuart, W. F. & Green, C. A. *Nature* **273**, 646–649 (1978).
2. Walker, A. D. M., Greenwald, R. A., Korth, A. & Kremser, G. *J. geophys. Res.* **87**, 9135–9146 (1982).
3. Allan, W., Poulter, E. M. & Nielsen, E. *Planet. Space Sci.* **31**, 1279–1289 (1983).
4. Jacobs, J. A. & Watanabe, T. *J. atmos. terr. Phys.* **25**, 267–279 (1963).
5. Woodman, R. F. & Hagfors, T. *J. geophys. Res.* **74**, 1205–1212 (1969).
6. Balsley, B. B. *J. atmos. terr. Phys.* **35**, 1035–1044 (1973).
7. Cap, F. F. *Handbook on Plasma Instabilities* Vol. 1, 351 (Academic, New York, 1976).
8. Brace, L. H., Theis, R. F. & Hoegy, W. R. *Geophys. Res. Lett.* **9**, 989–992 (1982).
9. Young, E. R., Burke, W. J., Rich, F. J. & Sagalyn, R. C. *J. geophys. Res.* **89**, 5565–5574 (1984).
10. Tanenbaum, B. S. & Mintzer, D. *Phys. Fluids* **5**, 1226–1237 (1962).
11. Iverson, I. B. *et al. J. geophys. Res.* **89**, 6775–6785 (1984).

Thickness of ice on perennially frozen lakes

Christopher P. McKay*, Gary D. Clow†, Robert A. Wharton Jr‡ & Steven W. Squyres‡

* Solar System Exploration Office, NASA-Ames Research Center, Moffet Field, California 94035, USA

† US Geological Survey, Menlo Park, California 94025, USA

‡ NASA-Ames Research Center, Moffet Field, California 94035, USA

The dry valleys of southern Victoria Land, constituting the largest ice-free expanse in the Antarctic, contain numerous lakes whose perennial ice cover is the cause of some unique physical and biological properties¹⁻³. Although the depth, temperature and salinity of the liquid water varies considerably from lake to lake, the thickness of the ice cover is remarkably consistent¹, ranging from 3.5 to 6 m, which is determined primarily by the balance between conduction of energy out of the ice and the release of latent heat at the ice-water interface and is also affected by the transmission and absorption of sunlight. In the steady state, the release of latent heat at the ice bottom is controlled by ablation from the ice surface. Here we present a simple energy-balance model, using the measured ablation rate of 30 cm yr⁻¹, which can explain the observed ice thickness.

The thermodynamics of ice formation have been studied extensively, particularly in application to the onset of freezing in waters seasonally covered with ice⁴ and the conditions of arctic sea ice⁵. The role of ablation in the formation of ice cover has been studied for arctic lakes⁶ and suggested for the Antarctic⁷. The antarctic lakes present a simplified physical problem for several reasons: (1) the thick perennial ice cover allows averaging of heat sources and sinks over the year; (2) that the lakes are confined allows a simple quantification of the incoming latent and sensible heat carried by liquid water; (3) the virtual absence of snowfall simplifies the treatment of the water mass balance, light penetration and heat conduction at the upper boundary.

The heat budget for the ice cover of a lake is shown schematically in Fig. 1. If the lake is large, lateral conduction of heat from the shore is negligible. Heat is added to the lake by geothermal heat flow from the bottom and by the influx of glacial meltwater whose latent heat is released at the ice-water interface when it freezes to form new ice. But as the meltwater is almost at the freezing point when it enters the lake⁸, the sensible heat input is small. Conduction of heat through the ice is the only means of heat exchange between the liquid water and the surface; hence, the rate of energy supply below the ice and the atmospheric heat flux must separately equal the conductive heat flow.

Although the input of solar energy, the amount of heat exchanged with the atmosphere and the inflow of meltwater all change markedly with season, the thick ice cover effectively averages these heat sources, a conclusion supported by the small ablation rate^{9,10}, 30 cm yr⁻¹, which implies a maximum variation in seasonal ice thickness of ~10%. The validity of neglecting seasonal variations is also supported by the observation that the ice thickness is larger than the *e*-folding depth of the annual temperature wave¹¹. Measurements of the thermal profile of the lakes also suggest that a steady state exists which is characterized by the annually-averaged conditions².

The balance of upward conduction of heat at a given depth in the ice cover and energy input below that level in the ice, assuming steady state conditions and averaging over the year, gives

$$k \frac{dT}{dz} = S(z) + L + F_g \quad (1)$$

where k is the thermal conductivity of ice, dT/dz is the gradient of annual mean temperature (T) with depth (z) in the ice, $S(z)$

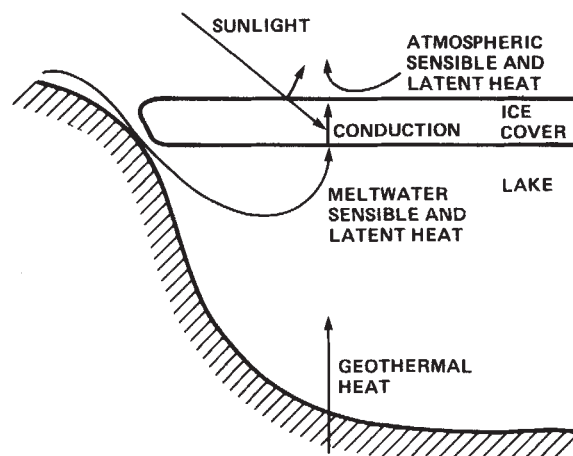


Fig. 1 Energy sources and sinks for perennially ice-covered lakes.

is the annual mean flux of solar energy absorbed below z and F_g is the geothermal heat flux. The heat flux due to latent heat release at the ice-water interface (L) is $L = vpl$ where v is the rate of formation of new ice averaged over the entire year, equal to the ablation rate, ρ is the density of the ice and l is the latent heat of fusion of water. To solve equation (1) for the temperature of the ice cover as a function of depth, we write $S(z) = (1-a)(1-r)S_0 \exp(-z/h)$ where a is the albedo of the lake, r is the fraction of the lake that is covered by dark absorbing material such as sand and silt, S_0 is the solar radiation incident on the lake surface and the exponential term gives attenuation with depth into the ice with an extinction path length of h (defined to be the cosine of the effective solar zenith angle divided by the extinction coefficient). The thermal conductivity of ice is approximated by¹² $k = b/T - c$ where b and c are constants taken to be 780 W m⁻¹ and 0.615 W m⁻¹ K⁻¹ respectively. Making these substitutions into equation (1) and integrating from the ice-water interface to the surface gives

$$Z = \frac{b \ln(T_0/T_s) + c(T_s - T_0) - S_0(1-a)(1-r)h(1 - \exp(-Z/h))}{vpl + F_g} \quad (2)$$

where T_0 is the temperature of the ice-water interface, T_s is the yearly averaged temperature of the surface, both in K, and Z is the equilibrium thickness of the ice cover. Equation (2) is an implicit expression for the average yearly thickness of the ice cover. In the limit that the ice cover is much thicker than the extinction path length ($Z \gg h$) and in the approximation that the thermal conductivity is independent of temperature ($k = \text{constant}$) and the geothermal flux is small, this equation reduces to

$$Z = \frac{k(T_0 - T_s) - S_0(1-a)(1-r)h}{vpl} \quad (3)$$

We now consider values of the environmental quantities that appear in equation (2). The temperature of the ice-water interface is almost 0 °C because the upper waters of all the lakes have very low salinity¹. The yearly average surface temperature is known only for the Lake Vanda region¹³ and is -20.0 °C; measurements at Scott Base, 100 km east and near the coast, are similar¹⁴.

The annual mean solar flux incident on the surface of the lake, S_0 , has been measured at Lake Vanda¹³ and is 104 W m⁻² and direct measurements of albedo, a , over Lakes Vanda and Bonney yield values of 0.5 and 0.65 respectively⁸. The optical properties of the lake ice and water of Vanda and Bonney¹⁵ suggest that the albedo varies with season and is strongly dependent on solar zenith angle, with higher albedos at lower angles. This effect results from the inhomogeneous scattering exhibited by particulate or optically rough surfaces¹⁶ and would be expected to be important at the low solar zenith angles typical of

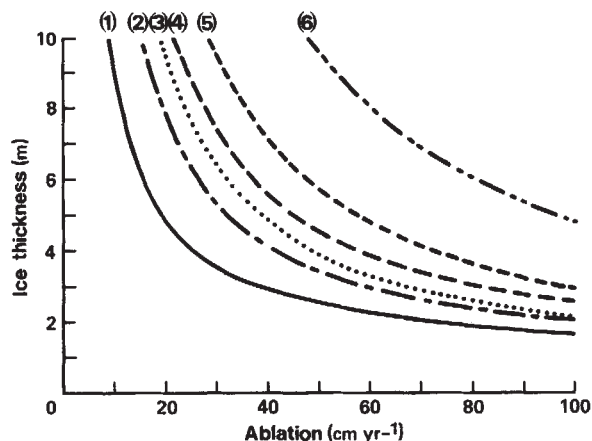


Fig. 2 Variation in ice-cover thickness as a function of ablation rate predicted by the model. The values of the parameters are as follows: $a = 0.6$, $F_g = 0.08 \text{ W m}^{-2}$, $S_0 = 104 \text{ W m}^{-2}$, and $r = 0.1$. The yearly-averaged surface temperature (T) and extinction path length (h) for the various curves are: line 1, $T = -20^\circ\text{C}$, $h = 1.0$; line 2, $T = -20^\circ\text{C}$, $h = 1.0$, $r = 0.25$; line 3, $T = -20^\circ\text{C}$, $h = 0.75$; line 4, $T = -25^\circ\text{C}$, $h = 1.0$; line 5, $T = -20^\circ\text{C}$, $h = 0.5$; line 6, $T = -20^\circ\text{C}$, $S_0 = 0$ (this is the ice cover that would result from no incident sunlight).

Antarctica. We have used an albedo of 0.6. The extinction path length of the ice cover (h) has been estimated indirectly from measurements of the total light level at the underside of the ice cover¹⁵. Extinction path lengths of 1 m and longer are reported^{8,15} and similar values can be inferred from measurements of the total transmission through the ice^{1,17}. The ice must be exceptionally clear as transmission data have been listed¹⁸ for ice and snow cover of lakes from which h values for clear ice in the range 0.6–1.3 m can be inferred, with even smaller values for 'bubbly' or 'snow' ice. Hence we use a nominal value of 1 m for h , but also calculate the ice-cover thickness for values of 0.75 and 0.5 m. No direct measurement of the areal extent of opaque material (r) on the different lakes has been attempted; it may vary considerably from lake to lake. We set r equal to 0.1, but also consider a case with $r = 0.25$.

The geothermal heat flow in the dry valleys has been measured at several sites¹⁹ as $\sim +0.08 \text{ W m}^{-2}$ and measurements of temperature gradients at the bottom of some of the lakes have given similarly small values^{2,20,21}.

Figure 2 shows the calculated equilibrium thickness of the ice cover plotted against the ablation rate for a range of the various values of parameters in equation (2) and also the ice thickness that would correspond to no incident sunlight ($S_0 = 0$). Ablation in the dry valleys is dominated by the dry downslope winds^{9,10,22} and may be approximately 30 cm yr^{-1} on all the lakes^{9,10}.

The model predictions are consistent with measurements^{1,10,23} of the ice-cover thickness, which in the dry valley lakes range from 3.6 to 6.5 m. One exception, Lake Vida, is frozen to its base²⁴, a depth of 11 m, which is explained by noting that the ice is anchored to the bottom and, as a result, ablation is not balanced by refreezing from below thereby eliminating this contribution to the heat flow through the ice cover. The effect of ablation in determining the thickness of the ice cover depends on the availability of liquid water beneath the ice as the source for new ice, and also on the latent heat released at the underside of the ice cover on freezing. The ablation thickness formula of Barnes⁶ applied to the antarctic lakes predicts a thickness of 15.8 m.

It is interesting to speculate on the ice cover that could have existed on possible former lakes in the equatorial regions of Mars. The lower parts of the Valles Marineris canyon system contain sedimentary deposits exhibiting rhythmic horizontal layering which is laterally continuous over large areas²⁵. The appearance of the sediments suggests deposition in standing water, a hypothesis consistent with other evidence for the action

of liquid water early in martian history. The suggestion²⁶ that lakes on Mars would freeze solid is tempered by our model, which indicates that ice thickness could be limited by ablation. Of course ablation rates for Mars are unknown, but a range of plausible values, of $1\text{--}10 \text{ cm yr}^{-1}$, is consistent with simple boundary layer theory and results in ice thicknesses ranging from 65 to 650 m. The above discussion is based on the present martian climate, in which equatorial temperatures²⁷ average -60°C , using an albedo of 0.5, $r = 0$; $S_0 = 180 \text{ W m}^{-2}$, $F_g = 0.0$ and $h = 1 \text{ m}$, and also assumes the presence of an aquifer from which martian lakes could be recharged²⁸. Relatively high ablation rates could result from downslope winds along the canyon walls; in addition temperatures in early martian history may have been considerably warmer than at present. More detailed evaluation of the hypothesis of martian palaeolakes may be made possible by careful study of the antarctic dry valley lakes.

C.P.M. and R.A.W. thank the Division of Polar Programs, NSF, for support of research in Antarctica. G.D.C. and S.W.S. thank the NASA Planetary Geology Program for support.

Received 24 September; accepted 22 November 1984.

1. Parker, B. C. *et al.* *J. Plankton Res.* **4**, 271–286 (1982).
2. Wilson, A. T. & Wellman, H. W. *Nature* **196**, 1171–1173 (1962).
3. Heywood, R. B. *Antarctic Freshwater Ecosystems: Review and Synthesis* (3rd SCAR symposium, Gulf, Houston, 1977).
4. Shaw, J. B. *Arctic* **18**, 123–132 (1965).
5. Baker, D. J. Jr., Radok, U. & Weller, G. *Rev. Geophys. Space Phys.* **18**, 525–543 (1980).
6. Barnes, D. F. *Air force surveys in geophysics* No. 129 (US Air Force, Cambridge, 1960).
7. Wilson, A. T. *Tuataara* **15**, 152–164 (1967).
8. Ragotzke, R. A. & Likens, G. E. *Limnol. Oceanogr.* **9**, 412–425 (1964).
9. Henderson, R. A. *et al.* *J. Glaciol.* **6**, 129–133 (1965).
10. Lyon, G. L. *N.Z. J. Geol. Geophys.* **22**, 281–284 (1979).
11. Campbell, G. S. *An Introduction to Environment Biophysics* (Springer, New York, 1977).
12. Wallace, D. & Sagan, C. *Icarus* **39**, 385–400 (1979).
13. Thompson, D. C., Bromley, A. M. & Craig, R. M. *N.Z. J. Geol. Geophys.* **14**, 477–483 (1971).
14. Thompson, D. C. & MacDonald, W. J. P. *Bull. N.Z. Dep. Sci. Ind. Res.* **140**, 37–56 (1961).
15. Goldman, C. R., Mason, D. T. & Hobbie, J. E. *Limnol. Oceanogr.* **12**, 295–310 (1967).
16. Squyres, S. W. & Veverka, J. *Icarus* **50**, 115–122 (1982).
17. Wharton, R. A. Jr., Parker, B. C. & Simmons, G. M. Jr. *Phycologia* **22**, 355–365 (1983).
18. Wetzel, R. G. *Limnology* (Saunders, New York, 1983).
19. Decker, E. R. & Bucher, G. J. *Antarctic J. U.S.* **12**, 102–104 (1977).
20. Shirlcliffe, T. G. L. *J. geophys. Res.* **69**, 5257–5286 (1964).
21. Bell, R. A. I. *N.Z. J. Geol. Geophys.* **10**, 540–546 (1967).
22. Bull, C. *Antarctic Res. Ser.* **9**, 177–194 (1966).
23. Parker, B. C. & Wharton, R. A. Jr. *Arch. Hydrobiol.* (in the press).
24. Catkin, P. E. & Bull, C. J. *Glaciol.* **6**, 833–836 (1967).
25. Lucchitta, B. K. & Ferguson, H. M. *J. geophys. Res. Suppl.* **88**, A553–A568 (1983).
26. Carr, M. H. *Icarus* **56**, 476–495 (1983).
27. Kieffer, H. H. *et al.* *J. geophys. Res.* **82**, 4249–4291 (1977).
28. Carr, M. H. *J. geophys. Res.* **84**, 2995–3007 (1979).

Hummocky cross-stratification is not produced purely under progressive gravity waves

Philip A. Allen

Department of Geology, University College, PO Box 78, Cardiff CF1 1XL, UK

Hummocky cross-stratification is a type of bedding in sedimentary rocks caused by the draping of sand around low-steepness three-dimensional bedforms. This distinctive form of stratification has been thought to be diagnostic of the former action of storm waves and hence its occurrence has been used to reconstruct ancient palaeogeographies and marine processes. Here I examine theoretical and empirical aspects of hummocky cross-stratification pertaining to its origin and argue that the notion that it was produced purely under progressive storm waves is incompatible with observed hummock spacings and possible values of the ratio a/λ under waves (a is the amplitude of oscillation near the bed and λ is the bedform spacing). My view is supported by recent observations on naturally-occurring hummocky megaripples from the sea floor of the inner Atlantic Shelf of North America.

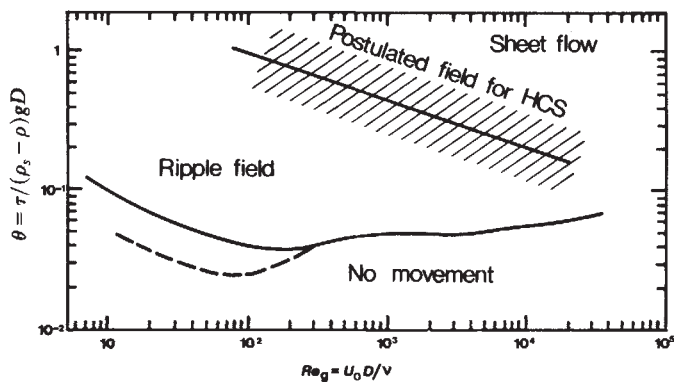


Fig. 1 The existence field for wave-formed ripples in the Shields plane, showing the postulated field for hummocky cross-stratification (HCS). Re_g is a grain Reynolds number.

Throughout the debate over the precise origin of 'truncated wave ripple laminae'^{1,2} and of its synonym 'hummocky cross-stratification'³, it has been consistently viewed as the most distinctive single sedimentary structure attributable to storms. The structure consists of an erosional lower contact overlain by low-angle undulating laminae than thin over hummock crests and thicken into intervening swales. Hummocks are generally 0.1–0.5 m high and spaced from one to several metres apart⁴. Hummocky cross-stratification is ubiquitous in sedimentary sequences known to be of shelf origin, such as the deposits of the Cretaceous seaway of the Rocky Mountain region, and most recorded examples occur in ancient winter storm zones⁵.

Some authors^{6–11} have postulated that hummocky cross-stratification originates under unidirectional highly concentrated suspensions generated during storms by (1) bypassing of the littoral zone by river-derived floods¹² and (2) offshore return of storm surges¹³, whereas others have favoured an origin under vigorous purely oscillatory flows generated by storm waves^{3,4,14,15}. Although these hypotheses are difficult to test because of the inaccessibility of the sea bed of continental shelves, investigations of the sea bed of the inner Atlantic Shelf of North America¹⁶ suggest that hummocky cross-stratification is a response to an along-coast geostrophic flow with wave orbital motions superimposed.

If hummocky cross-stratification is formed under intense oscillatory flows before the onset of bed planation⁴, the dimensions of the hummocks should be related to the flow conditions and oscillatory flow theory should be able to provide possible existence fields for these bedforms. The lower limit of the existence field of wave-formed ripples^{17,18} is defined by a relative stress similar to that used¹⁹ for unidirectional flows (Fig. 1), with shear stress estimated using a wave friction factor²⁰. The upper limit of ripple existence was shown to be

$$\theta_{cr} = 4.40(U_m D / \nu)^{-1/3} \quad (1)$$

using essentially a single set of data²¹; this is also in broad agreement with data from oscillating bed experiments²².

The planation of sand beds under waves may be explained by a hindered bedload settling effect²³. An approximate explicit form of the threshold for bed planation is

$$a\omega = 2.8(\gamma g)^{0.64} D^{0.22} \nu^{0.07} \omega^{-0.35} \quad (2)$$

where a is the near-bed amplitude of oscillation, ω is the wave frequency ($2\pi/T$), γ is the submerged specific gravity $(\rho_s/\rho) - 1$, where ρ_s is the sediment density and ρ is the fluid density, D is the sediment grain size and ν is the kinematic viscosity. Equation (2) can be expressed in terms of three dimensionless groups: $A = \gamma g D^3 / \nu^2$, which is a buoyancy index; $\Phi = (a\omega)^2 / \gamma g D$, a measure of sediment movement by waves; and a/D , which is related to grain roughness. Using these dimensionless groups

$$\Phi = 4.6(a/D)^{0.52} A^{-0.052} \quad (3)$$

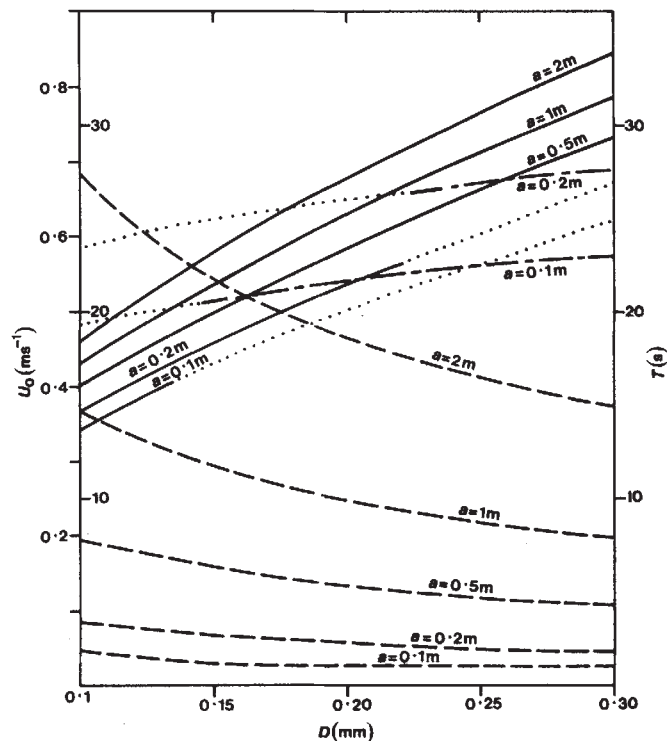


Fig. 2 Relationship between orbital velocity (U_0), wave period (T), near-bed amplitude of oscillation (a) and grain size (D) at the bed planation threshold. Bed planation velocities (left-hand ordinate) calculated using a Bagnold-type relative stress for $a/D > 1,000$ (—) and a hindered bedload settling equation for $a/D < 1,000$ (---); —, critical wave periods (right-hand ordinate). Curves for bed planation velocity are dotted (····) where a/D is inappropriate.

indicating the relatively weak influence of A on bed planation; using experimental values^{21,24,25} for A , this can be further simplified for the very fine to fine sands usually considered, that is, for $a/D < 10^3$, giving $\Phi = 4(a/D)^{0.52}$.

Hallermeier's hindered bedload settling criterion predicts very well the threshold velocity for planation found experimentally^{21,24,25} and applies to high-frequency waves occurring naturally in lakes or restricted seas.

A more appropriate and generally applicable threshold equation for lower frequency waves on oceanic coasts is based on a Bagnold-type relative stress, $\theta = \tau_0 / \gamma g D$. Using the wave friction factor, f_w , it can be shown that

$$\theta = \frac{1}{2} f_w (a^2 \omega^2 / \gamma g D) \quad (4)$$

For an almost planar bed, flow is likely to be transitional between laminar and turbulent²⁶, but adopting the analytical form for a smooth boundary²³ and regrouping into the dimensionless parameters, Φ , A and a/D , we have for the bed planation condition

$$\Phi = 27.4 \theta A^{0.1} (a/D)^{0.2} \quad (5)$$

Bagnold's^{27,28} theoretical studies suggest that θ at the transition to plane beds should be in the region of 0.5 for grain sizes of < 0.3 mm.

Suitable values of A and θ_{cr} can be substituted into equation (5) to give the simplest possible expression for sediment movement by wave energy for very fine and fine sands, $14.5(a/D)^{0.2} < \Phi < 22.4(a/D)^{0.2}$. As $\Phi = (a\omega)^2 / \gamma g D$ and $a\omega = U_0$, the bed planation velocity and critical wave period can be plotted against grain size for different near-bed amplitudes of oscillation (Fig. 2).

The foregoing discussion provides the basis for evaluating the near-bed processes responsible for the planation of a fine-sand bed under waves. The final constraint is the dimensions

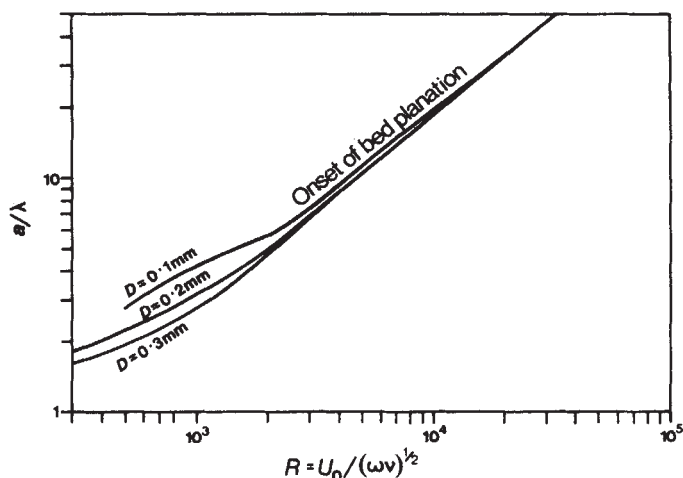


Fig. 3 Best-fit curves for low-steepness bedforms showing the variation of a/λ with wave Reynolds number, R , for various values of the sediment grain size, D .

of the hummocks, which are spaced from one to a few metres apart^{4,16} in the North American Cretaceous sediments. The near-bed amplitude of oscillation required to produce high-steepness vortex ripples of wavelength λ is²⁹ $a = 0.77\lambda$, the upper limit of orbital diameters being determined by grain size. The minimum value of a/λ for vortex ripple formation is³⁰ ~ 0.4 . These relationships are inappropriate for high wave Reynolds numbers (large amplitude of oscillation, long wave periods) where a/λ increases dramatically. The best-fit curve for low-steepness bedforms, based on the data of Sleath³¹ and Manohar²¹, has the form

$$a/\lambda = F(\beta D, R, \frac{\rho_s}{\rho}, \frac{(\rho_s - \rho)gD}{\rho\omega\nu}) \quad (6)$$

where $\beta D = (\omega/2\nu)^{1/2} D$ is a measure of the grain size relative to the viscous layer thickness and $R = U_0/(\omega\nu)^{1/2}$ is a form of wave Reynolds number. It can be expressed as

$$a/\lambda = 0.44((\rho_s - \rho)gD/\rho\omega\nu)^{0.098}(\beta D)^{-0.71}(\rho_s/\rho)^{-0.158}R^{0.070} \quad (7)$$

for $X < 3,000$ and

$$a/\lambda = 0.03((\rho_s - \rho)gD/\rho\omega\nu)^{0.432}(\beta D)^{-0.885}(\rho_s/\rho)^{-0.147}R^{0.056} \quad (8)$$

for $X > 3,000$, where

$$X = ((\rho_s - \rho)gD/\rho\omega\nu)(\rho_s/\rho)^{0.033}(\beta D)^{-0.523}R^{-0.042}.$$

It can be seen that the four groups on the right-hand side of equations (7) and (8) are dimensionless variables, in fact identical to those used by others³². The relationship between a/λ and wave Reynolds number, R , is shown in Fig. 3, using equations (3), (5), (7) and (8) and assumed values for the fluid and sediment densities and the kinematic viscosity. That a/λ is independent of grain size at high Reynolds numbers is clear.

Using these relationships, Fig. 4 shows that the predicted crest spacing of low-steepness bedforms is far smaller than the observed spacing of hummocks. I therefore propose that surface gravity waves alone cannot produce hummocky bedforms on the sea floor and what little data is presently available on natural occurrences of hummocky megaripples¹⁶ supports this conclusion. However, surface gravity waves may play an important part in the formation of hummocky bedforms in fine sand by first causing an enhancement of bed shear stress relative to that of a unidirectional current alone³³ and second, by introducing a strongly three-dimensional structure to near-bed water motion, thereby preventing substantial bedform migration.

Storm waves are always accompanied by steady flows such as coastal jets³⁴ or geostrophic currents³⁵. The solution of the

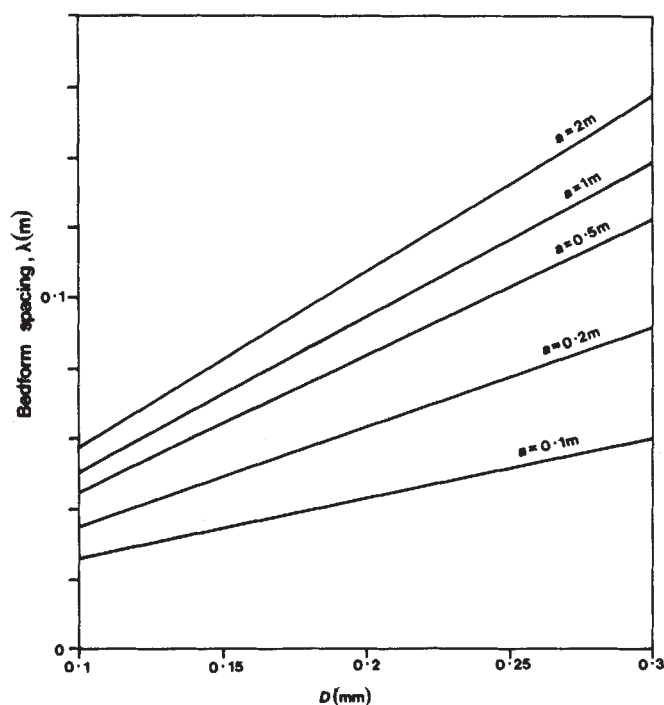


Fig. 4 Predicted crest spacing of low-steepness bedforms (λ) as a function of grain size (D) and near-bed amplitude of oscillation (a).

hummocky cross-stratification problem, and indeed of all storm sedimentation problems, must surely come from a detailed investigation of combined wave-current flows. In addition, a fuller understanding of the behaviour of the ratio a/λ under large amplitudes of oscillation and long wave periods is urgently required if progress is to be made in the palaeohydraulic interpretation of submarine sedimentary bedforms.

This letter owes much to the work of Robert Hallermeier (US Army Coastal Engineering Research Centre) and John Sleath (University of Cambridge), both of whom I thank.

Received 3 September; accepted 3 December 1984.

- Campbell, C. V. *J. sedim. Petrol.* **36**, 825-828 (1966).
- Campbell, C. V. *J. sedim. Petrol.* **43**, 765-778 (1971).
- Harms, J. C., Southard, J. B., Spearing, D. R. & Walker, R. G. *Soc. Econ. Paleont. Miner. Short Course Notes* **2** (1975).
- Dott, R. H. & Bourgeois, J. *Bull. geol. Soc. Am.* **93**, 663-680 (1982).
- Marsaglia, K. M. & Klein, G. de V. *J. Geol.* **91**, 117-142 (1983).
- Swagor, N. S., Oliver, T. A. & Johnson, B. A. in *Can. Soc. Petrol. Geol. Core Conf.*, Calgary, 78-95 (1976).
- Hamblin, A. P., Duke, W. L. & Walker, R. G. *Bull. Am. Ass. Petrol. Geol.* **63**, 460-461 (1979).
- Walker, R. G. *Geosci. Can. Rep. Ser.* **1**, 75-89 (1979).
- Wright, M. E. & Walker, R. G. *Can. J. Earth Sci.* **18**, 795-809 (1981).
- Mount, J. F. *J. sedim. Petrol.* **52**, 941-958 (1982).
- Brenchley, P. J. & Newall, G. *J. sedim. Petrol.* **47**, 819-833 (1977).
- Brenner, R. L. *Bull. Am. Ass. Petrol. Geol.* **64**, 1223-1244 (1980).
- Hayes, M. O. *Univ. Texas Bur. Econ. Geol. Rep. Invest.* **61**, 1-56 (1967).
- Bourgeois, J. *J. sedim. Petrol.* **50**, 681-702 (1980).
- Hunter, R. E. & Clifton, H. E. *J. sedim. Petrol.* **52**, 127-144 (1982).
- Swift, D. J. P., Figueiredo, A. G. Jr, Freeland, G. L. & Oertel, G. F. *J. sedim. Petrol.* **53**, 1295-1317 (1983).
- Komar, P. D. & Miller, M. C. *J. sedim. Petrol.* **45**, 697-703 (1975).
- Allen, J. R. L. *J. geol. Soc. Lond.* **136**, 673-682 (1979).
- Shields, A. *Versuchsanst. für Wasserbau und Schiffbau*, Berlin, 26 (1936).
- Jonsson, I. G. *Proc. 10th Conf. Coastal Engng I*, 127-148 (1967).
- Manohar, M. *Tech. Memo Beach Eros. Bd U.S. Corps Engrs*, 75 (1955).
- Bagnold, R. A. *Proc. R. Soc. A* **187**, 1-15 (1946).
- Hallermeier, R. *J. Nature* **297**, 53-55 (1982).
- Chan, K. W., Baird, M. H. I. & Round, G. F. *Proc. R. Soc. A* **330**, 537-559 (1972).
- Kennedy, J. F. & Falcon, M. M. *J. Hydr. Div. Am. Soc. Civ. Engrs* **91**, 1-55 (1965).
- Kamphuis, J. W. *J. Watways Harb. Div. Am. Soc. Civ. Engrs* **101** (WW2), 135-144 (1975).
- Bagnold, R. A. *Phil. Trans. R. Soc. A* **249**, 235-297 (1956).
- Bagnold, R. A. *Prof. Pap. U.S. geol. Surv.* **422-1** (1966).
- Miller, M. C. & Komar, P. D. *J. sedim. Petrol.* **50**, 173-182 (1981).
- Sleath, J. F. A. *Sea Bed Mechanics* (Wiley-Interscience, New York, 1984).
- Sleath, J. F. A. *J. Hydraul. Res.* **14**, 69-81 (1975).
- Mogridge, G. R. & Kamphuis, J. W. *Proc. 13th Conf. Coastal Engng, Vancouver*, 1123-1142 (1972).
- Bijker, E. W. *Delft Hydraul. Lab. Rep.* **50**, 1-142 (1967).
- Csanady, G. T. in *Lakes: Chemistry, Geology, Physics*, 21-64 (Springer, Berlin, 1978).
- Swift, D. J. P. & Rice, D. D. *Soc. econ. Paleont. Miner. spec. Publ.* **34**, 43-62 (1984).

Cd and Mn in the Alboran Sea and adjacent North Atlantic: geochemical implications for the Mediterranean

P. J. Statham*, J. D. Burton* & D. J. Hydes†

* Department of Oceanography, The University, Southampton SO9 5NH, UK

† Institute of Oceanographic Sciences, Wormley, Godalming, Surrey GU8 5UB, UK

Water flowing from the Mediterranean at the Strait of Gibraltar differs in oceanographic characteristics from inflowing Atlantic Ocean water because of the influence of climatic conditions in the Mediterranean Basin¹. Little is known about how geochemical processes affect concentrations of trace metals during this transformation because there are very few accurate published data (other than some recent values² for surface waters) for trace metals in the Mediterranean Sea, even though these are also required in assessing the impact of pollutant inputs. Here we present new data for the concentrations of dissolved Cd and Mn in the eastern Alboran Sea and the adjacent Atlantic Ocean and use them to estimate the flux of these metals through the Strait of Gibraltar. Although the Mediterranean is a source of Cd for the Atlantic Ocean, it acts as a sink for Mn entering in the inflowing Atlantic water.

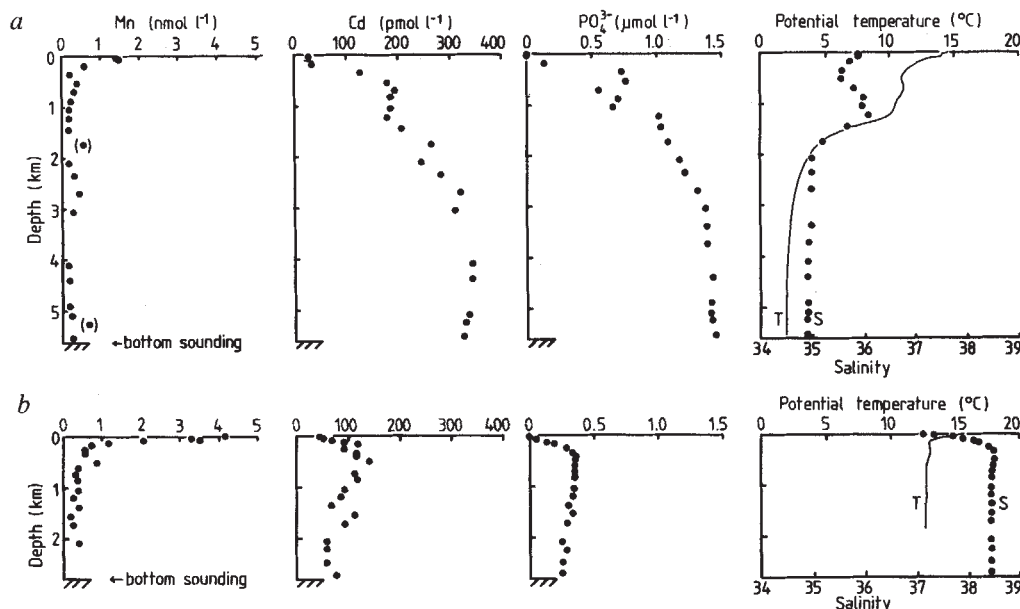
Samples were collected at Discovery Stations 10704 (North Atlantic Ocean; 39° 22' N, 12° 51' W; 9 April 1983) and 10707 (Eastern Alboran Sea; 37° 33' N, 00° 47' E; 17–18 April 1983), using rosette-mounted Teflon-coated 2.5-litre Go-Flo bottles, deployed with a conductivity-temperature-depth (CTD) probe to provide real-time data on the water column structure. After the samples were pressure-filtered from the bottles through 0.4-µm Nuclepore filters, trace metals were separated and concentrated at sea by extraction of the pyrrolidinedithiocarbamate and diethyldithiocarbamate complexes into 1,1,2-trichloro-1,2,2-trifluoroethane (Freon TF) and back extraction with nitric acid. The concentrates were analysed subsequently ashore by graphite-furnace atomic-absorption spectrophotometry. Analytical separations and critical manipulations were performed under clean conditions in a Class 100 laminar-flow hood. Full details of the procedure are given elsewhere³.

The vertical profile of dissolved Cd in the Atlantic Ocean (Fig. 1) shows a distribution consistent with other data^{4–8} with a general nutrient-like increase in concentration with depth. There is, however, a pronounced negative anomaly in the concentration of dissolved Cd, associated with the presence of nutrient-depleted Mediterranean water, shown by salinity and temperature data to influence the depth range 500–1,500 m. In the Mediterranean profile, Cd concentrations are lower, with much less pronounced vertical changes, similar to (or somewhat lower than) recently reported values⁹. In common with other oceanic regions¹⁰, a strong correlation between dissolved Cd and PO_4^{3-} has been observed for Atlantic waters^{4–8}, and has been interpreted in terms of related mechanisms of vertical particulate transport and cycling for these elements. In Fig. 2 the concentrations of dissolved Cd are plotted against those of PO_4^{3-} and NO_3^- for Stations 10704 and 10707 with corresponding data⁴ for a profile in the Sargasso Sea.

In the top 150 m of the profile in the Alboran Sea (that is, the waters overlying the Levantine Intermediate Water) the ratios $\text{Cd}/\text{PO}_4^{3-}$ and Cd/NO_3^- are higher than those characteristic of the deeper and Atlantic waters (Fig. 2). Higher surface concentrations of Cd in the Mediterranean relative to those in the adjacent Atlantic Ocean have been observed previously². These findings suggest an input of Cd to surface Mediterranean waters. The $\text{Cd}/\text{PO}_4^{3-}$ ratios in the deeper Alboran Sea are generally similar to those for the Atlantic Ocean. Profiles for other parts of the Mediterranean, particularly for the regions of intermediate and deep water formation, are needed to interpret the processes giving rise to the difference between surface and deeper waters, as are data on inputs to the system.

The flows through the Strait of Gibraltar provide the major import and export terms for the Mediterranean Sea¹¹. As the particulate fractions of the total concentrations of Cd¹² and Mn¹³ in open ocean waters are usually insignificant, a knowledge of the flow rates and dissolved trace metal contents of water masses provides a basis for simple mass-balance calculations². Spivack *et al.*², using this approach, but working from surface-concentration data alone, estimated that there was a net export of $2 \times 10^6 \text{ mol Cd yr}^{-1}$ from the Mediterranean. Water fluxes through the Strait of Gibraltar are $3.5 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$ for the surface inflow of Atlantic Ocean water and $3.3 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$ for the outflow of Mediterranean Water¹¹. A major component of the outflow is provided by the Levantine Intermediate Water at about 200–600 m, but there is also a significant contribution

Fig. 1 Vertical profiles of dissolved Mn, dissolved Cd, reactive PO_4^{3-} , potential temperature and salinity at *a*, Station 10704 (Eastern North Atlantic Ocean), and *b*, Station 10707 (Alboran Sea). Points in parentheses are probably affected by contamination. The underlined point is below the analytical detection limit. Blanks for Cd and Mn were, respectively, 0.05 and 0.07 nmol l^{-1} and the corresponding detection limits, estimated as twice the variance of the blank, were 0.03 and 0.07 nmol l^{-1} . Coefficients of variation of <5% were obtained for Cd at 0.19 nmol l^{-1} and Mn at 3.24 nmol l^{-1} . The accuracy of the data is supported by the results of a recent intercalibration study²². Dissolved NO_3^- and reactive PO_4^{3-} were determined onboard ship using automated analytical techniques based on standard procedures and salinities were measured using a Guildline salinometer.



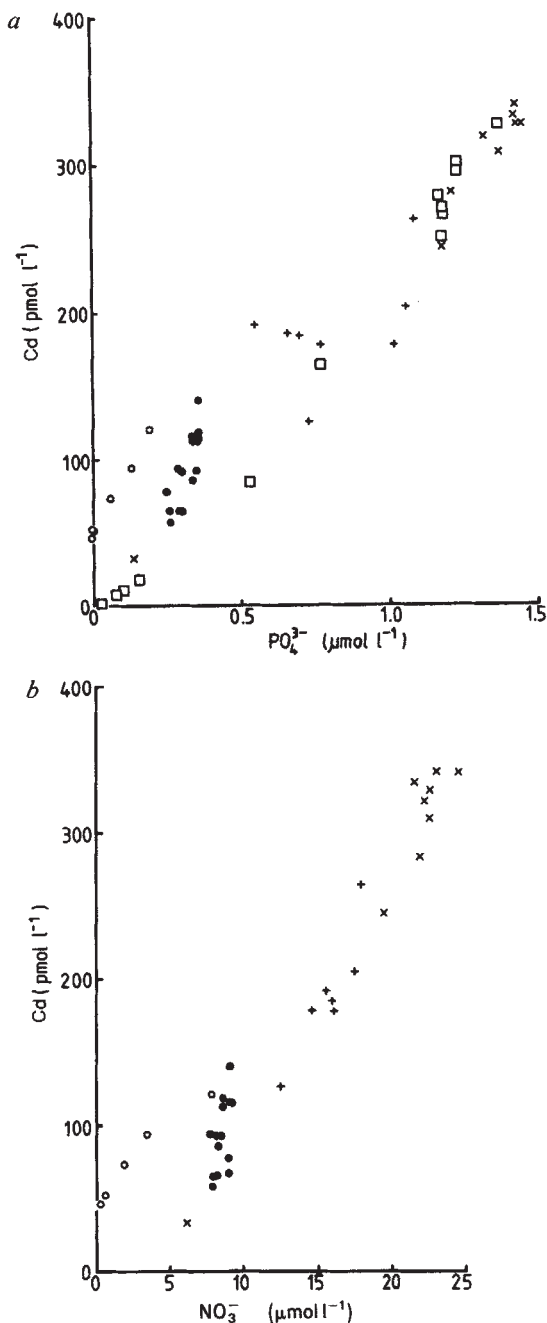


Fig. 2 Relationships between dissolved Cd and a, reactive PO_4^{3-} and b, NO_3^- for Stations 10704 and 10707. Mediterranean waters above 200 m ($\circ$); deeper Mediterranean waters ($\bullet$); Atlantic waters influenced by outflow of Mediterranean water (+); other Atlantic waters ($\times$). Data on Cd and PO_4^{3-} for a Sargasso Sea profile⁴ are also incorporated ($\square$).

from Mediterranean Deep Water¹⁴, ~25% of the total outflow. The data in Fig. 2 indicate that the respective concentrations of dissolved Cd in these waters are not greatly different (~110 and ~80 pmol l^{-1}); thus, the estimate of the total export of the element is not dependent critically on assumptions concerning the individual contributions of these two sources to the outflow. For present purposes, a concentration of 100 pmol l^{-1} is assumed. On this basis, and assuming the inflowing water to contain 19 pmol l^{-1} (see ref. 2), a net export of $2.6 \times 10^6 \text{ mol Cd yr}^{-1}$ is obtained.

River transport, mobilization from shelf sediments, human activities and atmospheric inputs have also been suggested as possible sources of Cd for the Mediterranean², but few data are available to evaluate their importance. Aeolian inputs of Cd to the western Mediterranean, however, appear high, estimated at

116 $\text{pmol cm}^{-2} \text{ yr}^{-1}$ (ref. 15). The net export of Cd suggested by the mass balance model corresponds to an average surface flux to the Mediterranean of 87 $\text{pmol cm}^{-2} \text{ yr}^{-1}$. Although there is little information on the extent and rate of dissolution in seawater of trace elements associated with aeolian particulates, measurements made using samples of airborne particles from the western coast of North America¹⁶ indicate that 80% or more of their Cd content rapidly enters solution in seawater. If a comparable fraction of the Cd in airborne particles introduced into Mediterranean surface water enters solution, then aeolian inputs could account for much of the supply of total dissolved Cd required by the mass balance model. The elevated surface Cd concentrations would be consistent with an aeolian source, although other processes, such as discrimination against Cd uptake under nutrient-limiting conditions¹⁷, might lead also to enhancement.

Aeolian inputs have been suggested as one source of Mn which could generate the surface enrichments observed for this element in the oceanic water column^{5,13,18}. Dissolved Mn at Stations 10704 and 10707 shows the same general distribution observed previously in the Atlantic Ocean^{4,5,19}, with high surface concentrations decreasing rapidly over the top few hundred metres to low and uniform deep water levels. Deeper water concentrations at both stations are similar at ~0.25 nmol l^{-1} , suggesting that scavenging and removal processes are operating similarly in both regimes, but the surface concentration in the Mediterranean (4.19 nmol l^{-1}) is much greater than that in the adjacent Atlantic Ocean (1.41 nmol l^{-1}). Earlier data⁹ show similar surface enrichment trends in Mediterranean waters, although the concentrations reported are generally higher than those given here.

A mass balance calculation for Mn in the Mediterranean has been performed in the same way as for Cd, using representative concentrations of 1.35 nmol l^{-1} for the inflowing surface water mass and 0.61 nmol l^{-1} for the outflowing water mass; it indicates a net input of Mn to the Mediterranean of $27 \times 10^6 \text{ mol yr}^{-1}$. Other possible inputs from the sources given above for Cd would increase the excess input to the Mediterranean even further. A very approximate estimate of the aeolian input of dissolved Mn to the Mediterranean has been made using flux data for the tropical North Atlantic Ocean¹⁵. If it is assumed that 50% of the Mn in the aeolian material enters solution¹⁶, a flux of $19 \times 10^6 \text{ mol Mn yr}^{-1}$ is obtained. Such an input is consistent with the elevated surface concentrations observed. The role of the Mediterranean as a sink for Mn is consistent with the short residence times reported for this element, in the range <1–25 yr^{19-21} for surface oceanic waters, relative to the mean residence time of ~100 yr for the Mediterranean waters themselves.

We thank the officers, crew and scientific complement of RRS *Discovery* on Cruise 137, and in particular Dr T. R. S. Wilson and T. J. P. Gwilliam. We also thank J. K. Clancy for assistance during sampling procedures. The work of P.J.S. and J.D.B. was supported by a NERC grant.

Received 22 August; accepted 21 November 1984.

1. Lacombe, H., Gascard, J. C., Gonella, J. & Béthoux, J. P. *Oceanol. Acta* **4**, 247–255 (1981).
2. Spivack, A. J., Huested, S. S. & Boyle, E. A. in *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E. A., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 505–512 (Plenum, New York, 1983).
3. Statham, P. J. *Analytica chim. Acta* (in the press).
4. Bruland, K. W. & Franks, R. P. in *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E. A., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 395–414 (Plenum, New York, 1983).
5. Burton, J. D., Maher, W. A. & Statham, P. J. in *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E. A., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 415–426 (Plenum, New York, 1983).
6. Ólafsson, J. in *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E. A., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 475–485 (Plenum, New York, 1983).
7. Yeats, P. A. & Campbell, J. A. *Mar. Chem.* **12**, 43–58 (1983).
8. Nürnberg, H. W., Mart, L., Rützel, H. & Sipos, L. *Chem. Geol.* **40**, 97–116 (1983).
9. Kremling, K. & Petersen, H. *'Meteor' Forsch. Ergebn.* **23**, 5–14 (1981).
10. Bruland, K. W. in *Chemical Oceanography*, (eds Riley, J. P. & Chester, R.) **8**, 157–220 (Academic, London, 1983).
11. Tchernia, P. *Descriptive Regional Oceanography* (Pergamon, Oxford, 1980).
12. Knauer, G. A. & Martin, J. H. *J. mar. Res.* **39**, 65–76 (1981).
13. Landing, W. M. & Bruland, K. W. *Earth planet. Sci. Lett.* **49**, 45–56 (1980).
14. Bryden, H. L. & Stommel, H. M. *J. mar. Res.* **40** (suppl.) 55–71 (1982).
15. Buat-Ménard, P. in *Air–Sea Exchanges of Gases and Particulates* (eds Liss, P. S. & Stinn, W. G. N.) 455–532 (Reidel, Dordrecht, 1983).
16. Hodge, V., Johnson, S. R. & Goldberg, E. D. *Geochem. J.* **12**, 7–20 (1978).

17. Boyle, E. & Huested, S. in *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 379-394 (Plenum, New York, 1983).
18. Klinkhammer, G. P. & Bender, M. L. *Earth planet. Sci. Lett.* **46**, 361-384 (1980).
19. Bender, M. L., Klinkhammer, G. P. & Spencer, D. W. *Deep-Sea Res.* **24**, 799-812 (1977).
20. Martin, J. H. & Knauer, G. A. *Earth planet. Sci. Lett.* **51**, 266-274 (1980).
21. Martin, J. H. & Knauer, G. A. *Earth planet. Sci. Lett.* **67**, 35-47 (1984).
22. International Council for Exploration of the Sea, Rep. CM 1983/E24, submitted to the Marine Environment Quality Committee.

A mid-Mesozoic extension across Central Asia?

William Leith

Lamont-Doherty Geological Observatory of Columbia University,
Palisades, New York 10964, USA

Although relatively little is known about basin formation north of the Himalayas, recent stratigraphic studies of several sedimentary basins (Tarim, Dzhungaria, Tsaidam) have demonstrated a common pattern of subsidence in post-Triassic times¹. All these basins show rapid subsidence concurrent with the onset of Himalayan orogenesis, presumably reflecting both tectonic loading and increased sediment supply. Before that time, however, subsidence curves have the characteristic 'concave-up' shape of rifted-margin sedimentary environments². Although data from these basins are not sufficiently detailed to justify a more thorough investigation, the Tadjik Depression, a sedimentary basin comprising the southwestern portion of the Soviet republic of Tadjikistan (Fig. 1), contains a continuous and relatively well documented record of post-Jurassic sedimentation, analysis of which suggests deposition in a rifted-margin environment, on crust thinned by ~50%.

As part of a US-Soviet scientific exchange, I have studied the rocks now exposed in folds and thrusts in the Tadjik Depression, to the north and west of the western syntaxis of the Himalayan range. The subsidence that accompanied the deposition of these sediments was examined with the joint aims of interpreting the tectonic setting in which the preorogenic deposition occurred and mapping the preorogenic structure of the basin. Stratigraphic data from the excellent exposures along the lower arm of Nurek Reservoir, 5 km south east of the town of Nurek, were used to estimate the temporal pattern of tectonic subsidence (details of the stratigraphy and subsidence calculations are given elsewhere^{3,4}). Stratigraphic sections from areas throughout the eastern Depression were then used to determine the spatial patterns of deposition and subsidence.

Jurassic-Recent strata are exposed where the Vakhsh river cuts downward through the upper plate of the Ionakhsh thrust. The post-Jurassic section (5,765 m) is divided into three major lithic units, each composed of strata with a significant number of common lithological features, reflecting deposition in a relatively uniform environment. These are: (1) the Sanglok group (not formally named) of early Cretaceous age; (2) the Pulisanga group of late Cretaceous-Palaeogene age; (3) the Tutkaul group, Neogene (Fig. 2). The lithological characteristics of the members of each group are described elsewhere³ and only the requisite features of the stratigraphy are presented here. The base of the section is predominantly halite (Upper Jurassic) and is tectonically disturbed; its thickness is not known, but more than 1,500 m of evaporite have been drilled in a borehole 5 km west of Nurek dam⁵. Evaporites mark the base of all of the sections south and east of the Illiac fault (see Figs 1, 4) and salt is actively doming in the southeastern Tadjik Depression (ref. 4 and W.L. and D. W. Simpson, in preparation).

In the Nurek region, there are no channels, stratal truncations or other sedimentary structures which indicate erosion and no significant nonconformities have been demonstrated during mapping. The Nurek section also has the thickest pre-Neogene section measured in the eastern Depression, which, for the purposes of this preliminary analysis, is assumed to be complete. It is also inferred from lithofacies data that the entire pre-

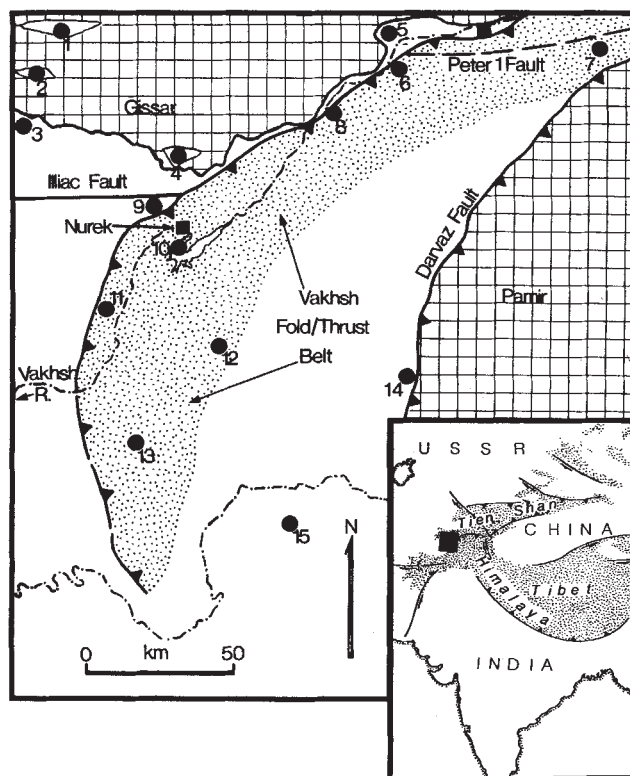


Fig. 1 Location map. The Tadjik Depression lies north and west of the north western end of the Himalayan arc (see inset). Stratigraphic sections (numbered dots) are shown. Sections 6-14 are from exposures in the Vakhsh fold/thrust belt; 1-5 from outliers of Depression sediments on the autochthon to the north. The detailed Nurek section (10) is located at Nurek reservoir, within the Vakhsh fold/thrust belt, just above the Ionakhsh thrust fault (structural details in ref. 24).

Neogene sequence was deposited in shoal water, including the Lower Cretaceous sandstones, interpreted as marine, which interfinger with shallow-water marine limestones in outcrops to the west of Nurek³.

Fossil assemblages have been described only from the rocks of the Pulisanga group, whose fauna are correlated biostratigraphically to the regionally-named suites³ (subgroups) and chronostratigraphically to the stages of the late Cretaceous and Palaeogene^{3,6,7}. The Sanglok group is dated by lithological correlation with the partly-fossiliferous marine sections to the west of Nurek noted above. Ages within the Tutkaul group are less certain (as they are based on lithological correlation with type section located 60 km to the north east), but I am not primarily concerned with these orogenic deposits.

The use of the backstripping method⁸ in the analysis of the stratigraphic sequence of an ancient rifted margin has been described in detail⁹. The significant assumptions used for the analysis of the Tadjik Depression are: (1) the basement-response function is calculated as an Airy-type isostatic compensation; (2) solution processes have not significantly altered the stratigraphic thicknesses (the section does not show extensive bedding-plane stylolitization or other evidence of solution cleavage); (3) the section is not completely compacted (instead, it is assumed that the rocks have a porosity typical of their lithology at their burial depths¹⁰).

Figure 3 presents the backstripping analysis as a 'tectonic' subsidence curve (shaded). This curve shows the approximate depth to which the basement would have subsided under water and in the absence of any sedimentary load (that is, because of some 'tectonic' process alone). For comparison, I also show 'thermal' subsidence curves, calculated from a one-dimensional lithospheric stretching model¹¹, for various amounts of thinning caused by rapid, constant-volume stretching of the lithosphere, followed by vertical heat conduction that results in cooling,

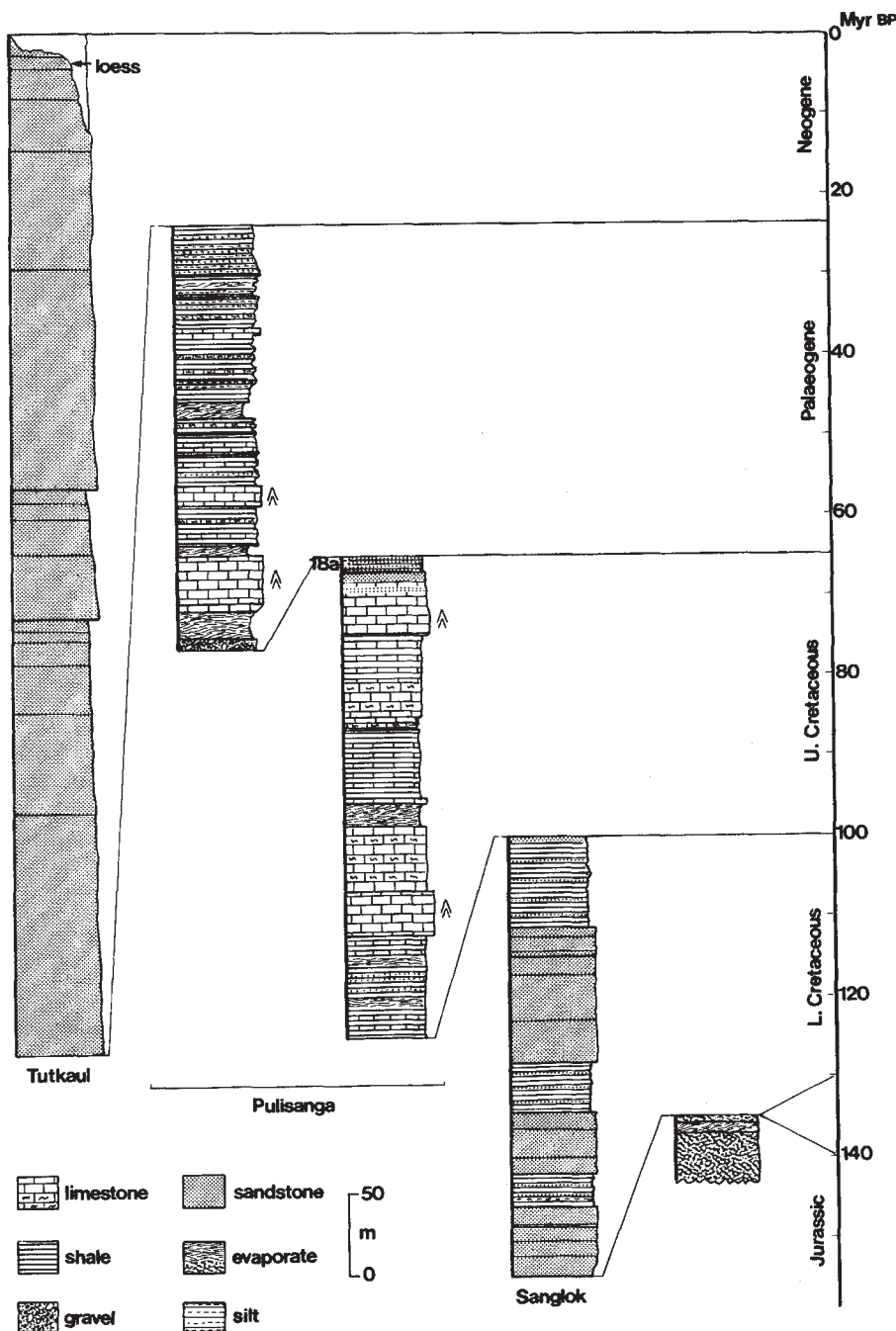


Fig. 2 Stratigraphic section for the Nurek reservoir area. The Sanglok group is composed predominantly of red marine sandstones and brown shales with minor siltstone beds. The sandstones are generally coarse and cross-bedded. In the upper portions of the group, isolated layers of calcareous sands are found and some silts show ripple marks. The Pulisanga group is composed primarily of interbedded limestone, shale, marl and gypsum, with uncommon beds of silt or sandstone and a single bed of gravel grading to conglomerate. Limestones are abundantly fossiliferous and, infrequently, bituminous. Green shales dominate the lower portion of the group, whereas red and brown shales are found in the upper portion. The Tutkaul group consists predominantly of red beds, predominantly a coarse clastic molasse that is relatively uncemented and very thickly bedded (in some parts >100 m). The boundaries between members are defined by breaks in the textural gradation that results from an increasing percentage of clays as one proceeds stratigraphically upwards within the sands of a single member. Parts of the group are probably aeolian.

thermal contraction and subsidence at the surface. The curve passing through the data is for a lithosphere stretched to twice its original length (stretching factor $\beta = 2$) and thinned to 50% ($= 1/\beta$) of its original thickness.

Between 130 and ~20 Myr BP, there is a close correspondence between the tectonic subsidence curve determined for the Pulisanga canyon section and the thermal subsidence curve (Fig. 3) in accordance with a correspondence noted between stretching-model subsidence curves and subsidence curves calculated by backstripping the stratal sequences of modern passive margins¹²⁻¹⁵. Thus, the present data are consistent with deposition on extended lithosphere, which supports the hypothesis¹⁶ that the Jurassic-Palaeogene strata of the Depression were deposited in an ancient rifted-margin tectonic environment. Since 20 Myr BP, the subsidence of the basin has increased, departing significantly from the thermal form, which presumably reflects both a mechanical loading and a dramatic increase in sediment supply and thereby coincides with orogenic uplift in the surrounding region.

Stratigraphic sections for 14 other sites in the eastern Tadjik

Depression^{17,18} (Fig. 1), are shown in Fig. 4, allowing a rough estimate of the distribution of subsidence over a large area. Sections 6-13 are all taken from outcrops of strata in overthrust plates within the Vakhsh fold/thrust belt, whereas sections 1-5 are taken from the region north of the northernmost thrusts of the fold/thrust belt (see Fig. 1). Section 14 is taken from the outcrop of strata in the thrust plate immediately adjacent to the Darvaz fault, which separates the Tadjik Depression sediments from the Upper Palaeozoic basement rocks of the Pamir.

There is a significant spatial variation in cumulative stratigraphic thickness by Eocene time, that is, before the initiation of fold/thrust deformation (marked by the onset of molassic sedimentation) and near the estimated time of the India-Asia collision¹⁹. In particular, there is an obvious discontinuity in pre-Eocene sediment thicknesses between the five northern sections and those to the south and east. These northern sections, outside the fold/thrust belt, also show significantly less (~1/3) subsidence than those located within the fold/thrust belt. In fact, two of these sections (4 and 5) are <25 km distant from sections (9 and 6) exhibiting more than twice the cumulative

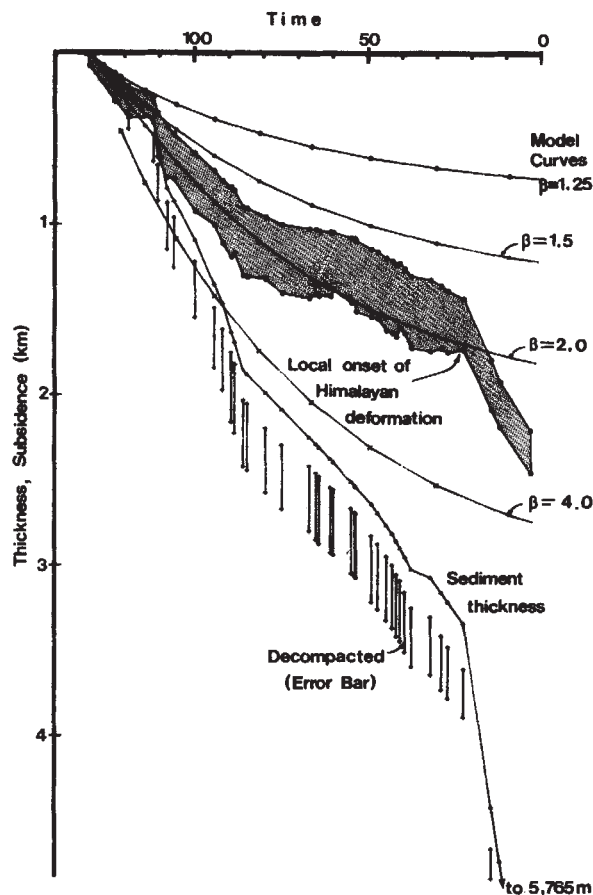


Fig. 3 Cumulative thickness and tectonic subsidence curve (shaded) for the Nurek section. Shown for comparison are the 'thermal' basement subsidence curves, predicted by a simple one-dimensional stretching model in which the lithosphere has been stretched by a factor (β), for several values of β . Data show typical post-extensional subsidence before the onset of Himalayan deformation. Note that that deformation occurs more than 15 Myr after the estimated time of collision between India and Asia (see ref. 19). Uncertainties in estimating the water depth at the time of deposition are included in the error bars.

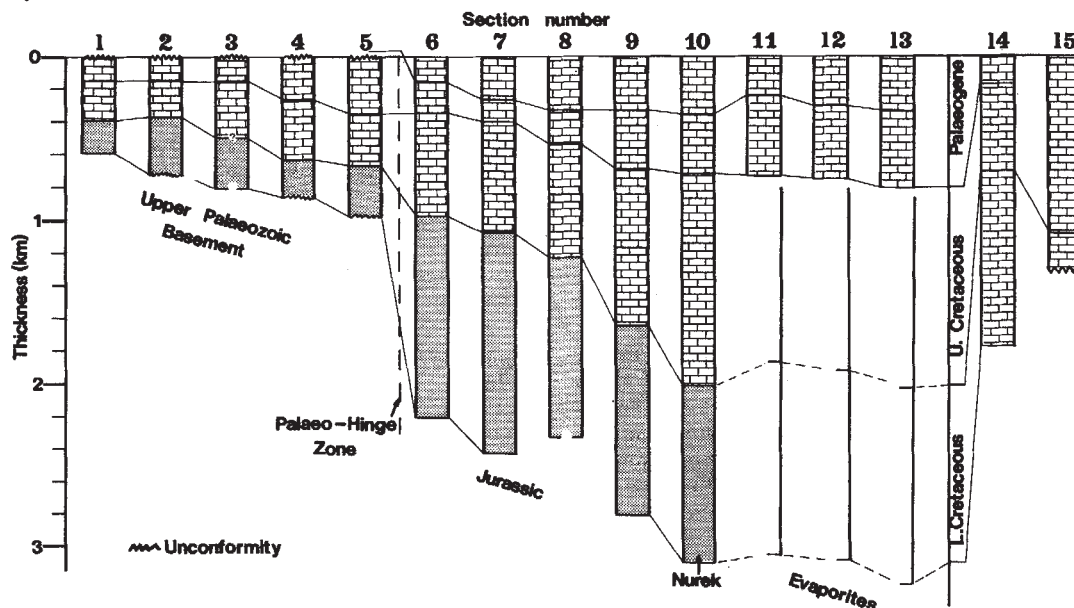


Fig. 4 Pre-Neogene stratigraphic sections for each of the 15 locations indicated in Fig. 1, arranged in order of increasing thickness. This arrangement is also approximately that from structurally external (left) to structurally internal (right), that is, the original relative north/south-location with thrust structure removed. Rocks of the Sanglok group (mostly sandstones) are in grey stipple; those of the Pulisanga group (mostly limestones and shales) are shown in a brick pattern. The major discontinuity in thicknesses between sections north of the Illiac and Peter-I-range faults (wherein post-Jurassic sections lie directly on Upper Palaeozoic basement without intervening Jurassic evaporites) and those south may indicate the approximate location of the palaeo-hinge zone of the ancient passive margin.

subsidence since Jurassic time. The large difference in thickness between these pairs of sections (4, 9 and 5, 6) is reflected stratigraphically by evidence for growth faulting; all of the rocks of corresponding chronostratigraphic units are thicker in section 9 than in section 4. But, all of the pre-Eocene rocks in both sections were probably deposited near sea-level.

Presently, sections 4 and 9 are separated by the Illiac fault, across which the crystalline crust to the south is displaced downward, as a block, by more than 2 km relative to the crust to the north²⁰. Thus, although the Illiac now slips in right-lateral motion²¹, it appears that most of the motion in the geological past was vertical slip by block-faulting, with the south side moving progressively downward. This allowed the accumulation of the thick sedimentary sequences deposited south of the fault.

Analyses of both modern and ancient rifted margins have demonstrated the existence of a common structural feature—the hinge zone¹³. The sediments of the coastal plain thicken up to the hinge zone, whereupon the basement falls away²²; the greatest subsidence and sediment accumulation is found seaward of the hinge zone and usually, at the hinge zone itself, the crust thins abruptly.

I suggest that the subsidence discontinuity in the northern Tadjik Depression is evidence for the mid-Mesozoic formation of a palaeo-hinge zone at or near the present Illiac fault. This hypothesis is supported by a seismic refraction profile²⁰, which showed that the crust thins, from 60–75 km in the Tien Shan to ~20–25 km in the Tadjik Depression, between the south slope of the Gissar range and the Illiac fault (V. I. Kulagin, personal communication). In addition, the presence of 20–25-km-thick crust beneath the sediments of the Tadjik Depression²⁰ is in accordance with the hypothesis that the lithosphere there has been thinned and extended by a factor of about two (that is, assuming a normal thickness for continental crust of ~40 km).

Although these observations support a previously-noted lithological similarity between the Jurassic–Oligocene rocks of the Tadjik Depression and those of modern passive margins¹⁶, they do not distinguish between sequences produced by intracratonic rifting and those produced by the rifting of a marginal arc during the formation of a marginal basin. Evidence to distinguish between these possibilities should lie within the syn-rift deposits, which are not exposed in the Tadjik Depression. Nevertheless, both possibilities are significant in that they point to a fundamental problem in Tethyan tectonics—the timing of suturing

along the northern margin of the Tethyan ocean. In either case, because the passive margin assemblage of the Tadjik Depression was not deformed until the late Tertiary (20–25 Myr BP), it appears that the Central Asian portion of the Asian margin was not completely sutured until well after the continental collision between India and 'Asia' (compare ref. 23). It seems that the northward motion of India continued to close oceanic basins that persisted in the late Tertiary, to the north of the terrains that were first accreted to its northern margin.

Received 17 August; accepted 22 November 1984.

1. *The Petroleum Potential of the Southeastern Asian Margin and China's Interior Basins* (Globex Consultant Services, Inc., 1982).
2. Dickinson, W. R. *Soc. Econ. Paleontol. Miner., spec. publ.* 22, 1–27 (1975).
3. Baratov, R. B. & Gaikili, V. N. *The Geology and Seismicity of the Nurek Hydroelectric Station District* (Academy of Science of Tadjikistan SSR., Dushanbe, 1962).
4. Leith, W. thesis, Columbia Univ. (1984).
5. Korovka, V. S. *et al.* in *Problems of the Petroleum Capacity of Tadjikistan* 4, 335–338 (Donish, Dushanbe, 1972).
6. Grossgeim, V. A. & Korobkov, I. A. *Paleogene System* (Nedra, Moscow, 1975).
7. Nalivkin, D. V. *Geology of the USSR* (Academy of Science of U.S.S.R., Moscow, 1962); (ed. Rast, N.) 514–533 (University of Toronto Press, 1973).
8. Watts, A. B. & Ryan, W. B. F. *Tectonophysics* 36, 24–44 (1976).
9. Bond, G. & Kominz, M. *Bull. geol. Soc. Am.* 95, 155–173 (1984).
10. Schmoker, J. W. & Halley, R. B. *Bull. Am. Ass. Petrol. Geol.* 66, 2561–2570 (1982).
11. McKenzie, D. *Earth planet Sci. Lett.* 40, 25–32 (1978).

This work was supported by contracts from the US Geological Survey, 14-08-0001-19743 and -21278. Fieldwork was supported by the Joint U.S.–U.S.S.R. Agreement in the Field of Environmental Protection (Area IX). I thank the following who made fundamental contributions to the work: W. Alvarez, A. M. Babaev, G. Bond, M. Kominz, L. M. Lyeskov, D. W. Simpson and A. Teremetski. I also thank Globex Consultant Services for allowing me to view their report. Lamont–Doherty contribution no. 3761.

12. Montadert, L. *et al.* Maurice Ewing Symp. Proc. Ser. 3, 154–186 (American Geophysics Union, Washington, D.C., 1979).
13. Watts, A. B. & Steckler, M. S. Maurice Ewing Symp. Proc. Ser. 3, 218–234 (American Geophysics Union, Washington, D.C., 1979).
14. Royden, L. & Keen, C. *Earth planet Sci. Lett.* 51, 343–361 (1980).
15. Keen, C. *et al.* *Oceanologica Acta* 26th Int. Geol. Cong. Proc., 123–128 (1981).
16. Leith, W. *Tectonics* 1, 303–318 (1982).
17. Gubin, I. Ye. *Patterns of Seismic Activity in Tadjikistan* (Academy of Science of USSR, Moscow, 1960).
18. Slavin, V. I. *Tectonics of Afghanistan* (Nedra, Moscow, 1976).
19. Molnar, P. & Tapponier, P. *Science* 189, 419–426 (1975).
20. Kulagin, V. I. & Kulagina, M. V. *Izvestiya Geol. Ser.* 1, 103–107 (1977).
21. Soboleva, O. V. *et al.* in *Searches for Earthquake Precursors in Prediction Polygons*, 65–70 (Nauka, Moscow, 1974).
22. Jansa, L. F. & Wade, J. A. *Can. Geol. Surv.* 74–30, 51–105 (1975).
23. Sengor, A. M. C. *Nature* 279, 590–593 (1979).
24. Leith, W. & Alvarez, W. *Bull. geol. Soc. Am.* (in the press).

Late Holocene deforestation and tree regeneration in the forest–tundra of Québec

Serge Payette* & Réjean Gagnon†

* Centre d'Études nordiques et Département de Phytologie, Université Laval, Sainte-Foy, Québec, Canada G1K 7P4

† Département des Sciences fondamentales, Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada G7H 2B1

Reconstructions of climate in the Holocene rely heavily on palaeoclimatic indicators such as altitudinal and latitudinal tree-line movements¹ inferred either from direct (macrofossil), or indirect (pollen), evidence of sites more or less distant from the present treelines^{2–8}. Long-term trends in tree regeneration on well-drained sites of the forest–tundra—the transition zone between the Boreal forest and the Arctic tundra zones—may also be used as ecological indicators of Holocene climatic changes. Charcoal found in soils of treeless or forest vegetation in the transition zone indicates respectively, failure or success in the post-fire regeneration of trees; as regeneration is influenced by climate, radiocarbon-dated charcoal can be used as a record of palaeoclimate. We suggest here that the widespread occurrence of treeless sites is the result of late Holocene deforestation involving climate–fire interactions and that disjunct lichen–forest sites are the outcome of successful regeneration sometime during the last 1,000 yr. This climatically induced process is acting at the site and the species levels, south of the present tree line.

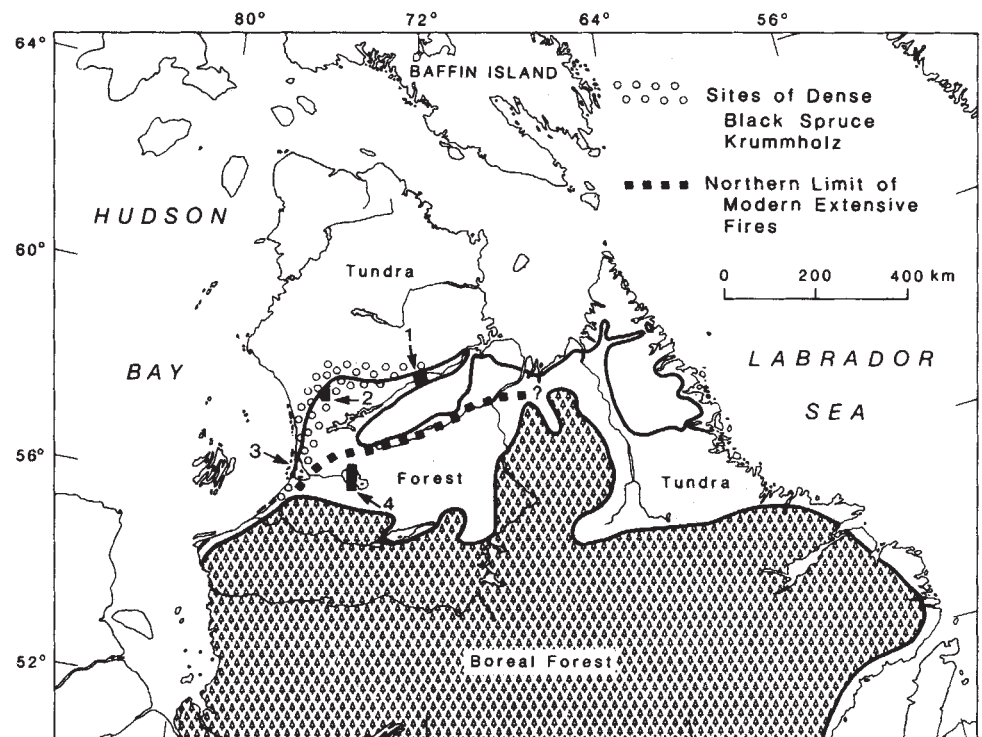
The northern Québec forest–tundra (Fig. 1) extends from the limit of continuous forest (Boreal forest) to the tree line⁹. Forests grow mainly into the lowlands in well-watered and protected sites, whereas treeless lichen-dominating vegetation is found on the interfluvies. Forests become increasingly scattered towards the north and occur as small groves near the latitudinal tree line, a similar situation to that outlined¹⁰ for the transition zone west of Hudson Bay. The occurrence of fires is widespread, especially in the more continental parts, and the presence of well-preserved coniferous charcoal, forming discrete layers at the top of mineral soils or buried into the soil profile by solifluction in treeless sites, suggests that coniferous vegetation was more extensive in the past. Charcoal found under the lichen ground cover, south of the modern tree line, has been interpreted as evidence of past movements of the north–central Canadian forest–tundra^{11,12}; this interpretation, assumed (1) that a finding

of charcoal is a direct proof of the presence of a former forest line and (2) that the spatio-temporal distribution of radiocarbon-dated charcoal shows a strong latitudinal shifting of the so-called forest line—defined¹³ as the northern edge of the adjoining Boreal forest, where forests cover 50% of the soil surface—and of Boreal forest zone, owing to climatic change. The underlying hypothesis rests on a conception of vegetation zones suggesting global movements of the forest–tundra, south and north of its modern position, in an oscillatory pattern, in response to climatic forcing. In our view, however, the problem is somewhat different, if one considers the forest–tundra as the product of a once densely-populated coniferous zone that experienced a long-term deforestation sometime after the regional postglacial climatic maximum.

Our study was conducted in four different regions of the forest–tundra in a south–north direction, from the limit of continuous forest to the tree line (Fig. 1). Radiocarbon-dated conifer charcoal from 116 well-drained sites was stratified on a habitat basis that reflects differential post-fire tree regeneration. The samples were collected systematically along transects at the sites, which are characterized by similar soil conditions supporting lichenic vegetation in treeless, krummholz (stunted spruces) and forest sites. Coniferous charcoal was identified from deciduous charcoal using standard criteria—anatomical characteristics such as lack of vessel elements, presence of tracheids, cone fragments and conifer leaves—but because coniferous charcoal at these sites does not indicate the exact growth forms that burned, it is often difficult to identify the composition of the pre-fire vegetation, unless wood remains are also present. Three conifer-fire chronologies, related to the last fire that occurred at each site of the treeless, krummholz and forest habitats, respectively, were constructed using weighted radiocarbon dates (Fig. 2).

In Fig. 2a, the data indicate that treeless vegetation was formed during the last 3,000 ¹⁴C-yr when coniferous tree species (black spruce, *Picea mariana* (and tamarack, *Larix laricina*) were extinguished because of a failure in sexual regeneration. These sites previously supported a sparse stunted black-spruce cover, occasionally with isolated tamarack trees, as suggested by various growth forms of wood remains. The former presence of white spruce (*Picea glauca*) in our sites seems unlikely, except in the Richmond Gulf area (Fig. 1) where the species extends throughout the maritime fog belt¹⁴, in a geographical position similar to that¹⁵ on the west coast of Hudson Bay; prismatic replacement of white spruce by black spruce, as suggested¹⁶ for Labrador, remains uncertain for the three continental regions (Fig. 1) and needs to be documented. The distribution of ¹⁴C-

Fig. 1 The northern Québec forest-tundra zone lying between the limit of continuous forest and the tree line, according to ref. 9. Sites of dense krummholz are typically positioned along the Hudson Bay seaboard and the tree line where fire is infrequent and of limited extent. Modern extensive fires, mapped from field surveys and aerial photographs, coincide with the geographical distribution of extensive forest vegetation mainly found in the southern part of the forest-tundra. Study areas: (1) Leaf River; (2) Bush Lake; (3) Richmond Gulf; (4) Clearwater Lake.



charcoal dates from dense black-spruce krummholz sites which have been free of fire disturbance for 400–600 yr and in some cases for more than 1,000 yr (Fig. 2b) is somewhat similar to that of Fig. 2a. These persistent communities are characterized by the presence of numerous wood remains lying at the ground surface. Sites from the third habitat are occupied by old-growth lichen-forest that regenerated around 900–800 BP and around 600–450 BP (Fig. 2c) and are unique in that they have been free of fire disturbance since these two fire periods. They usually exhibit well-preserved wood remains that reveal their long-term regeneration history.

The latitudinal position of the sampled charcoals (Fig. 3) shows that the onset of treeless communities was metachronous, although more fires were registered during the last millennium, and also that the long-term deforestation and conifer extinction are not related to distance from the present tree line (Fig. 1). This situation may be associated with a progressive incidence of areally smaller burns, characteristic of marginal forest-tundra and tundra sites¹⁷, reflecting an overall decrease in the supply of burning fuels.

The three conifer fire chronologies (Fig. 2) stress that differential post-fire tree regeneration occurred on well-drained sites of the forest-tundra during the last 3,000 ¹⁴C-yr. The post-fire communities now established on these sites are a direct response to a long-term regeneration process shown schematically in Fig. 4. The importance of post-fire tree regeneration at each site depended on the regeneration potential of the pre-fire communities, either forest (with trees bearing normal cones) or krummholz (with stunted spruces bearing only a few small cones, sometimes none); possibly on shorter or longer fire intervals^{18,19} and on the timing between a fire occurrence and the onset of cold or mild climatic conditions. (A similar scheme has also been outlined²⁰ for Holocene dune development in Québec.) From 3,000 BP to the present, the expansion of treeless vegetation (Fig. 2a) has primarily been the consequence of krummholz and forest burns, as suggested by wood remains with the former krummholz being possibly the outcome of ancient lichen woodland that failed to regenerate as trees because of the onset of cooling. Expansion of treeless vegetation through fire disturbance was somewhat decelerated around 2,100–1,800 BP, 1,050–800 BP, 650–450 BP and now, probably because of better spruce regeneration associated with the milder conditions pertaining

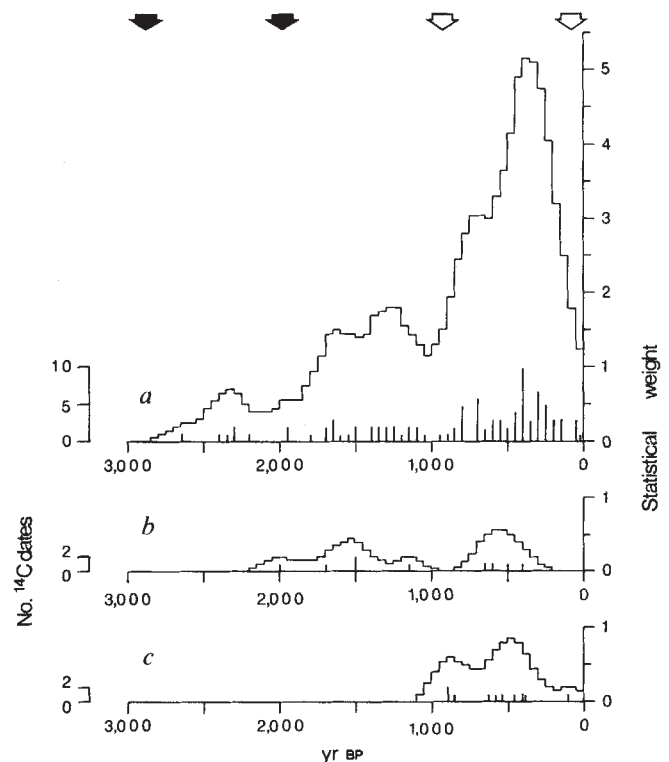


Fig. 2 Cumulative statistical-weight histogram of the 116 ¹⁴C dates from charcoal found in well-drained sites of the northern Québec forest-tundra zone. The histogram was constructed using methods given in ref. 26. The statistical weight is calculated for each 50-yr class interval (¹⁴C yr) from probability tables for a normal distribution ($\pm 2\sigma$). Also included is the distribution of individual ¹⁴C dates (vertical lines). ¹⁴C-charcoal dates of the last conifer fire that occurred in: a, present-day treeless sites; b, krummholz sites; and c, tree stands. Black arrows, macrofossil evidence of minor tree-line expansion north of the present tree line in the Leaf River area, according to ref. 3. White arrows, macrofossil and modern tree evidence of reforestation on well drained sites of the forest-tundra according to refs 3 and 14, in the Leaf River area and Richmond Gulf area, respectively.

Fig. 3 Distribution of ^{14}C -charcoal dates (solid lines, $\pm\sigma$) according to latitude. All samples from Leaf River (1) were collected at an altitude of 150–200 m above sea level; at 125–140 m in Bush Lake (2); 30 m in Richmond Gulf (3); and 235–300 m in Clearwater Lake (4). Treeless sites were often above and also at the same altitude as krummholz and old forest sites. $\circ$, Treeless site; $\square$, krummholz site; $\triangle$, forest site.

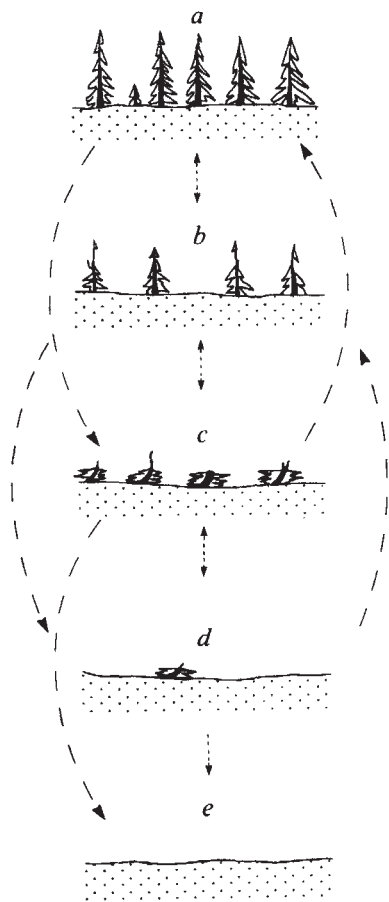
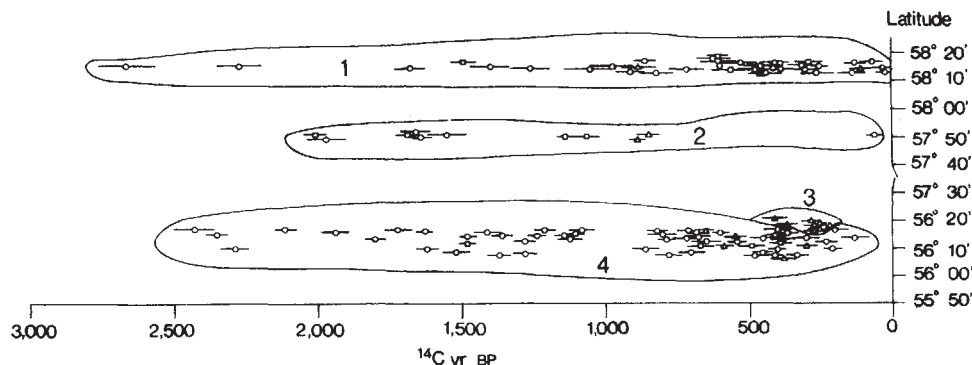


Fig. 4 The reforestation-deforestation process in the northern Québec forest-tundra zone (northern part) during the Holocene. Fire recurrence in spruce forest stands (a) during late Holocene cooling (downward arrows) caused a shift towards depauperate coniferous populations (b, c, d); subsequent fires culminated in total deforestation indicated by treeless communities (e). Short warming periods over the last 1,000 yr were responsible for local reforestation (upward arrows). Reforestation from long-distance seed dispersal occurred recently at some treeless sites¹⁴ near the tree line (not indicated here).

in forest stands around 900–800 BP and 600–450 BP. However, the forest sites are far less numerous than those recording deforestation and conifer extinction. Tamarack macrofossil evidence of minor tree-line expansion and local reforestation during these periods was also noted in the study area (Fig. 2). Finally, it is thought that most of the old-growth krummholz (Fig. 2b) represent forest remnants²¹, forming a transient stage towards treeless communities in the long-term regression of the coniferous cover at those latitudes (Fig. 4).

We conclude that the structure of the modern forest-tundra zone is the result of long-term tree regeneration dynamically related to Holocene climatic changes and to fire history. These ecological factors together contributed to progressive regression of the coniferous cover, forests and krummholz, respectively, and to local tree-species extinction, more particularly between 3,000 and 2,100 BP, 1,800 and 1,050 BP, 800 and 650 BP and 450 and ~100 BP. The spatio-temporal distribution of the radiocarbon-dated charcoals (Fig. 3) does not support any displacement of the transition zone in a way similar to that described for the north-central Canadian forest-tundra from charcoal data^{11,12} and from qualitative pollen data⁷. Instead of large-scale movements of the forest-tundra, our data suggest that this zone developed its modern facies at least during the last 3,000 ^{14}C -yr period of deforestation south of the present treeline, at a time when no important movements of the tree line were reported in western^{8,22,23}, central²⁴ and eastern^{3,25} Canada. Evidence of past tree-line displacements may be found north of the present tree line where krummholz stands grow (Fig. 1) in the area undisturbed by major fires for many centuries (Figs 1, 2). Therefore, their location gives the minimum magnitude for past tree-line movements related to black spruce. Because of the disjunct distribution of forest biomass and the infrequency of fires near the tree line, the northernmost part of the forest-tundra will probably be less susceptible to fluctuations than the southernmost part where frequent and large wildfires could increase the incidence of deforestation events.

We thank Dr Louise Filion (Québec) for helping to improve the manuscript and Mr Yvan Grenier for drawing the figures. This work has been supported by the Ministère de l'Éducation of Québec (FCAC Program) and by the Natural Sciences and Engineering Research Council of Canada.

Received 14 May; accepted 9 November 1984.

- Lamb, H. H. *Climatic History and the Future*, Vol. 2 (Methuen, London, 1977).
- Denton, G. H. & Karlén, W. *Quat. Res.* **7**, 63–111 (1977).
- Gagnon, R. & Payette, S. *Géogr. phys. Quat.* **35**, 52–72 (1981).
- Kearney, M. S. & Luckman, B. H. *Science* **221**, 261–263 (1984).
- Kullman, L. *Boreas* **9**, 101–106 (1980).
- LaMarche, V. C. *Quat. Res.* **3**, 632–660 (1973).
- Nichols, H. *Occ. Pap. 15* (University of Colorado, Institute of Arctic and Alpine Research, 1975).
- Ritchie, J. C. & Hare, F. K. *Quat. Res.* **1**, 331–342 (1971).
- Payette, S. *Nordicana* **47**, 3–23 (1983).
- Rowe, J. S. *Can. Dep. n. Affairs nat. Resources Branch Bull.* **123** (1959).
- Bryson, R. A., Irving, W. N. & Larsen, J. A. *Science* **147**, 46–48 (1965).
- Sorenson, C. J., Knox, J. C., Larsen, J. A. & Bryson, R. A. *Quat. Res.* **1**, 468–473 (1971).
- Larsen, J. A. *Ecol. Monogr.* **35**, 37–59 (1965).
- Payette, S. & Filion, L. *Can. J. for. Res.* (in the press).
- Ritchie, J. C. *Ecology* **38**, 429–435 (1957).
- Lamb, H. F. *J. Ecol.* **72**, 37–59 (1984).
- Wein, R. W. *Arctic* **29**, 213–222 (1976).
- Black, R. A. & Bliss, L. C. *Can. J. Bot.* **56**, 2020–2030 (1978).
- Black, R. A. & Bliss, L. C. *Ecol. Monogr.* **50**, 331–354 (1980).
- Filion, L. *Nature* **309**, 543–546 (1984).
- Nichols, H. *Arctic* **29**, 38–47 (1976).
- Ritchie, J. C. *Ecol. Monogr.* **47**, 401–423 (1977).
- Spear, R. W. *Nordicana* **47**, 61–72 (1983).
- Kay, P. A. *Quat. Res.* **11**, 125–140 (1979).
- Richard, P. *Paléophytogéographie postglacière en Engava par l'analyse pollinique*, Paléo-Québec **13** (1981).
- Ochietti, S. & Hillaire-Marcel, C. *Géogr. phys. Quat.* **31**, 123–133 (1977).

Ice-free cryopreservation of mouse embryos at -196°C by vitrification

W. F. Rall* & G. M. Fahy

American Red Cross, Cryobiology Laboratory,
9312 Old Georgetown Road, Bethesda,
Maryland 20814, USA

The failure of complex mammalian organs, such as the kidney, to function following freezing to low temperatures is thought to be due largely to mechanical disruption of the intercellular architecture by the formation of extracellular ice¹⁻⁵. Classical approaches to the avoidance of ice formation through the imposition of ultra-rapid cooling and warming rates⁶⁻⁸ or by gradual depression of the equilibrium freezing point during cooling to -80°C ⁹⁻¹³ have not been adequate. An alternative approach¹⁴⁻¹⁶ relies on the ability of highly concentrated aqueous solutions of cryoprotective agents to supercool to very low temperatures. At sufficiently low temperatures, these solutions become so viscous that they solidify without the formation of ice, a process termed vitrification. When embryo suspensions are cryopreserved using conventional procedures, this supercooling behaviour allows intracellular vitrification, even in the presence of extracellular ice¹⁷⁻²⁰. We have therefore used mouse embryos to examine the feasibility of obtaining high survival following vitrification of both the intra- and extracellular solutions and report here that in properly controlled conditions embryos seem to survive in high proportions after cryopreservation in the absence of ice.

Successful vitrification requires the use of a highly concentrated yet effectively non-toxic solution of cryoprotectants. Our vitrification solution (VS1) was based on solutions described elsewhere¹⁶ and consists of 20.5% w/v dimethyl sulphoxide (DMSO), 15.5% w/v acetamide, 10% w/v propylene glycol and 6% w/v poly(ethylene glycol) (PEG, relative molecular mass (M_r) 8,000) in a modified Dulbecco's saline (HB1) at pH 8.0. Similarly concentrated solutions of cryoprotectants are known to vitrify based on a variety of physical measurements^{16,21-23} as well as freeze-fracture microscopy⁸. VS1 (total concentration ~ 13 molal) has been shown to vitrify when cooled at rates above $5^{\circ}\text{C min}^{-1}$ both by differential scanning calorimetry (D. S. Reid, unpublished data) and by the procedure^{16,24} of direct visual inspection.

Cryoprotectant toxicity and osmotic injury are influenced by the concentration of cryoprotectant and the time and temperature of exposure. We controlled these factors as follows. Embryos were first transferred to an HB1 solution containing one-quarter of the cryoprotectant levels in VS1 (25% VS1) for 15 min at $\sim 20^{\circ}\text{C}$. The embryos shrank and then re-expanded to their initial volumes, indicating complete permeation of DMSO, acetamide and propylene glycol. The embryos were then exposed to more concentrated solutions in a cold room ($\sim 4^{\circ}\text{C}$) to reduce toxicity. Embryos were observed to shrink (dehydrate) and remain shrunken when exposed to these concentrated solutions at 4°C , indicating no further cryoprotectant permeation at this temperature. However, dehydration concentrates both intracellular protein and the previously introduced intracellular cryoprotectants and should thereby permit intracellular vitrification without additional cryoprotectant permeation^{16,20}. The short times required for dehydration also minimize toxicity. Dehydration has the further advantage of allowing cells to tolerate rapid dilution of the cryoprotectants after cryopreservation and warming¹⁶.

A high proportion of embryos pre-equilibrated in 25% VS1 as described above survived (based on development to expanded blastocysts within 48 h of culture) after exposure at 4°C to 75% VS1 for 40 min (94.3% survival, $n = 88$), 87.5% VS1 for

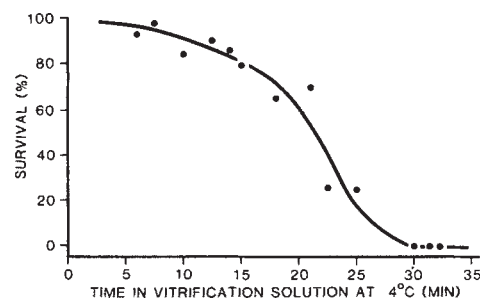


Fig. 1 Survival of eight-cell mouse embryos as a function of the time of exposure to the vitrification solution (VS1). Embryos were obtained from immature N: NIH(S) females that were superovulated and mated with N: NIH(S) males. Embryos were collected and handled as described previously²⁹. The embryos were washed in HB1 solution, a modified Dulbecco's saline³⁰ containing 1 mM phosphate and supplemented with 20 mM HEPES buffer, 0.33 mM sodium pyruvate, 5.56 mM glucose, 3 mg ml⁻¹ bovine serum albumin and 100 IU ml⁻¹ penicillin G. Groups of 14–25 embryos were placed in a 1:4 dilution of VS1 (25% VS1 = 5.125% w/v DMSO, 3.875% acetamide, 2.5% propylene glycol and 1.5% PEG in HB1) for 15 min at 20°C to allow complete permeation of the cryoprotectants. The embryo suspensions were then placed in a cold room ($\sim 4^{\circ}\text{C}$) where the embryos were exposed to the vitrification solution in two steps. First, the embryos were transferred to a 1:2 dilution of VS1 (50% VS1 = 10.25% w/v DMSO, 7.75% acetamide, 5% propylene glycol and 3% PEG in HB1) for 10 min and then finally into VS1. After various periods of exposure to VS1, embryos were immediately transferred into 50% VS1 for 10 min and then into 25% VS1 for 10 min at $\sim 4^{\circ}\text{C}$. The embryo suspension was then warmed to $\sim 20^{\circ}\text{C}$ and the remaining cryoprotectants were removed in two equal steps. Finally, the embryos were washed in PB1 solution³¹ and cultured in Brinster's BMOC-3 medium³². Survival was measured as the percentage of cultured embryos developing to the expanded blastocyst stage within 48 h. Each data point represents an average of 30 embryos (range 14–56).

20 min (94.7% survival, $n = 224$) or 90% VS1 for 20 min (93.6% survival, $n = 115$). Exposure to VS1, however, produces time-dependent injury (Fig. 1). High survival (80%+) is observed after exposure for up to 15 min, but survival decreases rapidly after longer exposure times. This decrease in survival could, however, be prevented at lower temperatures. For example, embryos exposed to VS1 for 10 min at $\sim 4^{\circ}\text{C}$ will tolerate an additional 60 min of exposure if the temperature is reduced to -20°C .

Following osmotic equilibration in VS1 or in various dilutions of VS1, embryos and their suspending medium were cooled either by plunging into liquid nitrogen or by being placed in a gradually cooling bath ($20^{\circ}\text{C min}^{-1}$). At the higher concentrations—100, 90 and 85% VS1—embryos survived in high proportions after both rapid and slow cooling (Fig. 2). However, at the lower concentrations—75 and 75% VS1—only about half the embryos survived after rapid cooling, and none survived after slow cooling.

The effect of warming rate on embryo survival is shown in Table 1. The temperature of the embryos was not permitted to exceed 4°C during either slow or rapid warming. Embryos warmed at $300^{\circ}\text{C min}^{-1}$ or above showed high survival independent of cooling rate, whereas embryos warmed at $10^{\circ}\text{C min}^{-1}$ were always killed. This result is presumably associated with devitrification (crystallization) at warming rates below $300^{\circ}\text{C min}^{-1}$. During warming at $10^{\circ}\text{C min}^{-1}$, devitrification was readily visible as a whitening of the suspending medium.

We propose that the high survival observed when embryos are cooled in 100% VS1 is associated with the absence of ice formation both extracellularly and intracellularly. Evidence for extracellular vitrification was described above. Intracellular vitrification is implied by earlier studies in which embryos were frozen slowly in 1.5 M DMSO²⁰. Slow freezing of 1.5 M DMSO

* Present address: Rio Vista International, Route 9, Box 242, San Antonio, Texas 78227, USA.

Table 1 Effect of warming rate on the survival of eight-cell mouse embryos cooled to -196°C

Container	Cooling rate ($^{\circ}\text{C min}^{-1}$)	Warming rate ($^{\circ}\text{C min}^{-1}$)	Total no. of embryos (replicate samples)	Survival (%)	s.e.m.
Embryos cooled in VS1					
Ministraw	2,500	2,500	343 (13)	87.8	1.9
Ministraw	20	2,500	110 (4)	83.6	3.9
Glass tube	500	300	171 (6)	80.7	4.0
Ministraw	2,500	10	63 (3)	0.0	—
Embryos cooled in 90% VS1					
Ministraw	2,500	2,500	188 (7)	84.6	3.7
Ministraw	20	2,500	204 (8)	80.8	4.9
Glass tube	500	300	151 (5)	87.4	2.2
Ministraw	2,500	10	68 (3)	0.0	—

Embryo suspensions were prepared and cryopreserved as described in Fig. 2 legend with two exceptions. First, some embryos were placed in a 10×75 -mm glass test tube containing 0.2 ml of the appropriate vitrification solution. Second, some ministraws were warmed slowly ($\sim 10^{\circ}\text{C min}^{-1}$) by transfer into a 16×150 -mm glass tube containing ethanol precooled to -110°C and allowed to warm in room temperature air. Slowly-warmed straws were removed from the alcohol at $\sim -5^{\circ}\text{C}$, placed in an ice-water bath and then brought into the cold room for immediate embryo recovery and dilution. The rates of rapid cooling and warming for straws are estimated from unpublished data of S. P. Leibo and the lower rates were measured using thermocouples. 90% VS1 solution was prepared by mixing 9 parts of VS1 and 1 part HB1: 18.45% w/v DMSO, 13.95% acetamide, 9% propylene glycol and 5.4% PEG in HB1.

to -40°C yields a concentrated residual solution that will vitrify when cooled to $< -130^{\circ}\text{C}$. Light microscopical observations indicate that embryos dehydrate when exposed to these solutions during slow freezing to -40°C and will undergo intracellular vitrification on subsequent rapid cooling¹⁸⁻²⁰. Therefore, embryos dehydrated to a similar extent by exposure to VS1 at

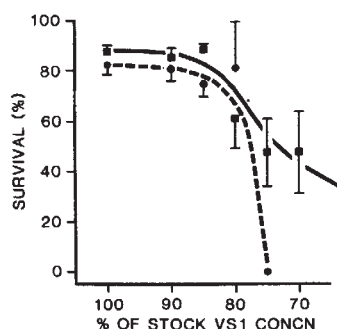


Fig. 2 Survival of eight-cell mouse embryos cooled to -196°C and warmed rapidly, as a function of the total concentration of cryoprotectants in the suspending solution. Embryos were collected and exposed to the 25% VS1 solution at $\sim 20^{\circ}\text{C}$ and the 50% VS1 solution at $\sim 4^{\circ}\text{C}$ as described in Fig. 1 legend. Groups of 15–40 embryos were then placed in one of several solutions containing 70–100% of the stock VS1 concentration. These solutions were prepared by diluting VS1 with HB1. After dehydration in these solutions (< 5 min), the embryos were pipetted into the centre of a $45\text{-}\mu\text{l}$ drop of the same suspending solution in a 0.25-cm^3 plastic insemination straw (IMV, L'Aigle, France). Ten minutes after transfer of the embryos to VS1, the heat-sealed straws were cooled either rapidly ($2,500^{\circ}\text{C min}^{-1}$) by direct immersion in liquid nitrogen (-196°C) or 'slowly' ($\sim 20^{\circ}\text{C min}^{-1}$) to -130°C in a steadily cooling isopentane bath before transfer to liquid nitrogen. After a storage period of 1–30 days, straws were warmed rapidly by vigorous shaking in a 0°C water bath ($2,500^{\circ}\text{C min}^{-1}$). Immediately after warming, the contents of the straw were expelled into a Petri dish containing 3 ml of the 50% VS1 solution at $\sim 4^{\circ}\text{C}$ and the embryos diluted and cultured as described in Fig. 1 legend. Approximately 98% of the cooled embryos were recovered and cultured. Values are means $\pm$ s.e.m. The number of replicate samples and the total number of embryos recovered after thawing (in parentheses) were, in order of decreasing concentration for rapidly cooled samples (■), 13 (343), 7 (188), 1 (21), 2 (52), 7 (223), 6 (225) and 4 (142), and for slowly cooled samples (●), in the same order, 4 (110), 8 (204), 2 (41), 2 (42) and 2 (38).

4°C would be expected to undergo intracellular vitrification on subsequent cooling. Our proposal is supported by the dependence of survival on the warming rate (Table 1). Embryos are killed by slow warming, a condition which would permit both extra- and intracellular devitrification (crystallization). Definitive proof of intracellular vitrification after cooling in VS1, however, will require further studies.

The survival of embryos cooled in 90 and 85% VS1 is more difficult to explain because ice will form in these solutions, especially during slow cooling. One speculative explanation for the unexpectedly high survival is related to the ability of a variety of biological and synthetic macromolecules to facilitate vitrification at relatively low concentrations^{15,16,25}. Cellular dehydration of embryos at 4°C will cause the endogenous intracellular macromolecules (for example, proteins) to concentrate and may therefore allow the intracellular solutions to vitrify more readily than the suspending medium. The death of embryos cooled in 75 and 70% VS1 presumably results from intracellular freezing during cooling and/or warming due to insufficient concentrations of intracellular solute.

Our results demonstrate that at temperatures as high as 4°C , mouse embryos can tolerate exposure to and osmotic dehydration in a concentrated solution of cryoprotectants which is capable of vitrifying. Moreover, embryos in VS1 survive cooling to -196°C and rapid warming in conditions which lead to vitrification of the extracellular and, presumably, the intracellular solutions. Recent experiments have demonstrated that human monocytes retain full functional integrity after exposure to and cooling in 93% VS1²⁶, suggesting that this approach to cryopreservation is applicable to other types of cells and tissues. Finally, current efforts to cryopreserve organs by vitrification are supported by evidence that mouse embryos in VS1 readily survive vitrification at moderate cooling rates and at warming rates which may well become feasible for large specimens^{27,28}, and by the absence of time-dependent cryoprotectant toxicity at -20°C and by the possibility that culture assays may allow the repair of cryoprotectant-induced injury observed¹⁶ using immediate tests of viability.

We thank P. M. Schmidt for supplying mated, superovulated mice and Dr H. T. Meryman for encouragement. This work was supported by grants GM 17959 and BSRG 2 507 RR05737 from NIH and the American Red Cross. This is contribution 644 from the American Red Cross, Blood Services Laboratories.

Note added in proof: Normal late-stage fetuses and live offspring result from the transfer of vitrified embryos to foster mothers (W.F.R., M. J. Wood and C. Kirby, in preparation).

Received 14 September; accepted 12 November 1984.

1. Pegg, D. E., Jacobsen, I. A., Armitage, W. J. & Taylor, M. J. in *Organ Preservation* Vol. 2 (eds Pegg, D. E. & Jacobsen, I. A.) 132-144 (Churchill Livingstone, Edinburgh, 1979).
2. Fahy, G. M. *Cryobiology* 18, 550-570 (1981).
3. Pegg, D. E. & Diaper, M. P. in *Organ Preservation, Basic and Applied Aspects* (eds Pegg, D. E., Jacobsen, I. A. & Halasz, N. A.) 389-393 (MTP, Lancaster, 1982).
4. Hunt, C. J. *Cryobiology* 21, 385-402 (1984).
5. Jacobsen, I. A. & Pegg, D. E. *Cryobiology* 21, 377-384 (1984).
6. Luyet, B. J. *Biodynamica* 1, P. 29, 1-14 (1937).
7. Luyet, B. J. & Gehehen, P. M. *Life and Death at Low Temperatures* (Biodynamica, Normandy, 1940).
8. Rapatz, G. & Luyet, B. *Biodynamica* 10, 193-210 (1968).
9. Farrant, J. *Nature* 205, 1284-1287 (1965).
10. Elford, B. C. & Walter, C. A. *Nature* 236, 58-60 (1972).
11. Elford, B. C. & Walter, C. A. *Cryobiology* 9, 82-100 (1972).
12. Fahy, G. M. *Cryobiology* 17, 371-388 (1980).
13. Kemp, E., Clark, P. B., Anderson, C. K. & Parsons, F. M. *Proc. Eur. Dialysis Transplant Ass.* 3, 236-240 (1966).
14. Fahy, G. M. & Hirsch, A. in *Organ Preservation, Basic and Applied Aspects* (eds Pegg, D. E., Jacobsen, I. A. & Halasz, N. A.) 399-404 (MTP, Lancaster, 1982).
15. MacFarlane, D. R., Angell, C. A. & Fahy, G. M. *Cryo-Letters* 2, 353-358 (1981).
16. Fahy, G. M., MacFarlane, D. R., Angell, C. A. & Meryman, H. T. *Cryobiology* 21, 407-426 (1984).
17. Leibo, S. P., McGrath, J. J. & Cravalho, E. G. *Cryobiology* 15, 257-271 (1978).
18. Rall, W. F., Reid, D. S. & Farrant, J. *Nature* 286, 511-514 (1980).
19. Rall, W. F. in *Frozen Storage of Laboratory Animals* (ed. Zeilmaker, G. H.) 33-44 (Fischer, New York, 1981).
20. Rall, W. F., Reid, D. S. & Polge, C. *Cryobiology* 21, 106-121 (1984).
21. Boutron, P., Delage, D., Rousit, B. & Korber, C. *Cryobiology* 19, 550-564 (1982).
22. Rasmussen, D. & Luyet, B. *Biodynamica* 11, 33-44 (1970).
23. Rasmussen, D. H. & MacKenzie, A. P. *Nature* 220, 1315-1317 (1968).
24. Angell, C. A. & Sare, E. J. *Cryo-Letters* 1, 257-260 (1980).
25. MacKenzie, A. P. in *Water Relations of Foods* (ed. Duckworth, R. B.) 477-503 (Academic, London, 1975).
26. Takahashi, T. et al. *Cryobiology* (submitted).
27. Burdette, E. C. in *Organ Preservation for Transplantation* (eds Karow, A. M. Jr & Pegg, D. E.) 213-259 (Dekker, New York, 1981).
28. Burdette, E. C., Wiggins, S., Brown, R. & Karow, A. M. Jr *Cryobiology* 17, 393-402 (1980).
29. Rall, W. F., Mazur, P. & McGrath, J. J. *Biophys. J.* 41, 1-12 (1983).
30. Dulbecco, R. & Vogt, M. J. *exp. Med.* 99, 167-182 (1954).
31. Whittingham, D. G. *Nature* 233, 125-126 (1971).
32. Brinster, R. L. in *Growth, Nutrition, and Metabolism of Cells in Culture* (eds Rothblat, G. H. & Cristofalo, V. J.) 251-286 (Academic, New York, 1972).

Retinotopic order appears before ocular separation in developing visual pathways

Glen Jeffery*

Department of Pharmacological and Physiological Sciences,
University of Chicago, 947 E. 58th Street,
Chicago, Illinois 60637, USA

In mammals, the major subcortical visual structures receive projections from both eyes, with the uncrossed projection being smaller than the crossed. Each projection is arranged as a separate orderly map of one hemiretina. Although these hemiretinal maps are separate in the nuclei, they are aligned so that the representations of points in the visual field are in register, thus there is a continuity of visual field representation between them¹. During the early development of the binocular pathways, terminals from the two eyes overlap almost entirely. As development proceeds, terminals arising from each eye segregate to form the adult pattern²⁻⁵. In the present study, local retinal lesions were made in ferrets at various stages in development before the separation of the projections from the two eyes. A neuronal tracer was then injected into the damaged eye, defining the pattern of projection from that eye. As reported here, the lesion resulted in a limited interruption in the pattern of terminal label on both sides of the brain, demonstrating that terminals from each eye are arranged in an orderly retinotopic manner at this stage. Hence, during later development, as one projection is reduced relative to the other, the two maps must slide in relation to each other.

In the ferret (gestation period 41 days), retinal ganglion cell axons start to penetrate subcortical visual structures on embry-

onic day 28 (E28); by E36 they have invaded all of the lateral geniculate nucleus (LGN). On the day of birth (D0), retinal projections from both eyes occupy the whole nucleus. During the next 9 days, the two projections segregate progressively, with the ipsilateral projection retracting to part of the binocular region of the dorsal LGN which is located caudally. The contralateral projection withdraws from this part of the caudal nucleus over the same time period⁶. In the present study I have used pigmented and albino animals ranging from the late prenatal to the early postnatal period. This included the following pigmented animals: 2 E38, 7 E39, 6 E40, 15 D0, 10 D1, 8 D2, 4 D5, and 4 D9; and 5 D0 and 3 D2 albinos. The prenatal animals were delivered by caesarian section. Animals were anaesthetized by hypothermia, then the skin over the left eye was cut and the exposed eye rotated to reveal part of its posterior aspect. A small section of aluminium foil was placed against a part of the sclera, and the fine tip of a heated thermal probe placed briefly against the foil; this caused a small section of the eye to fuse to the foil and produced a localized retinal lesion. The lesion site varied between animals. The lesioned eye was then injected with the tracer horseradish peroxidase (HRP), and ~10 h later the animals were deeply anaesthetized and perfused with a solution containing 1.5% paraformaldehyde and 1.5% glutaraldehyde. The brains were removed and sectioned either horizontally or sagittally, and processed to reveal anterogradely transported tracer⁷. The sections were lightly counterstained with neutral red. The injected eye was opened to confirm that the retina had been lesioned. The size of the lesion varied between animals, but in the majority it occupied ~20% of the retinal area.

Independent of age, in every animal in which a localized retinal lesion could be identified, and in which there was successful anterograde transport of HRP, a limited interruption could be identified in the pattern of terminal label, in one or more central visual structures. In the dorsal LGN no terminal or fibre labelling could be identified within this interruption. The border between labelled and unlabelled regions in the contralateral projection was sharp (see Fig. 1a). There was no obvious difference in the precision of this border between animals at E40 or later, nor was there any obvious difference between the sharpness of this border in the dorsal LGN (Fig. 1a) compared with that in the superior colliculus (SC; Fig. 1d). When lesions were confined to the nasal retina, the interruption in the pattern of label in the contralateral dorsal LGN was confined to the rostral segment of the nucleus (Fig. 1b); in the adult animal this region becomes the monocular sector, receiving its projection from the peripheral nasal retina⁵. This lesion also produced an interruption in the pattern of terminal label in the contralateral caudal SC; in the adult this region would also be monocular. When lesions were confined to the temporal retina, the interruption in the pattern of terminal label in the contralateral dorsal LGN was confined to the caudal region (Fig. 1c); in the contralateral SC it was confined to the rostral region (Fig. 1d). Both of these regions in the adult animal would represent binocular sectors of the visual field⁵. There was no obvious difference between the pigmented and the albino animals in the order of this projection.

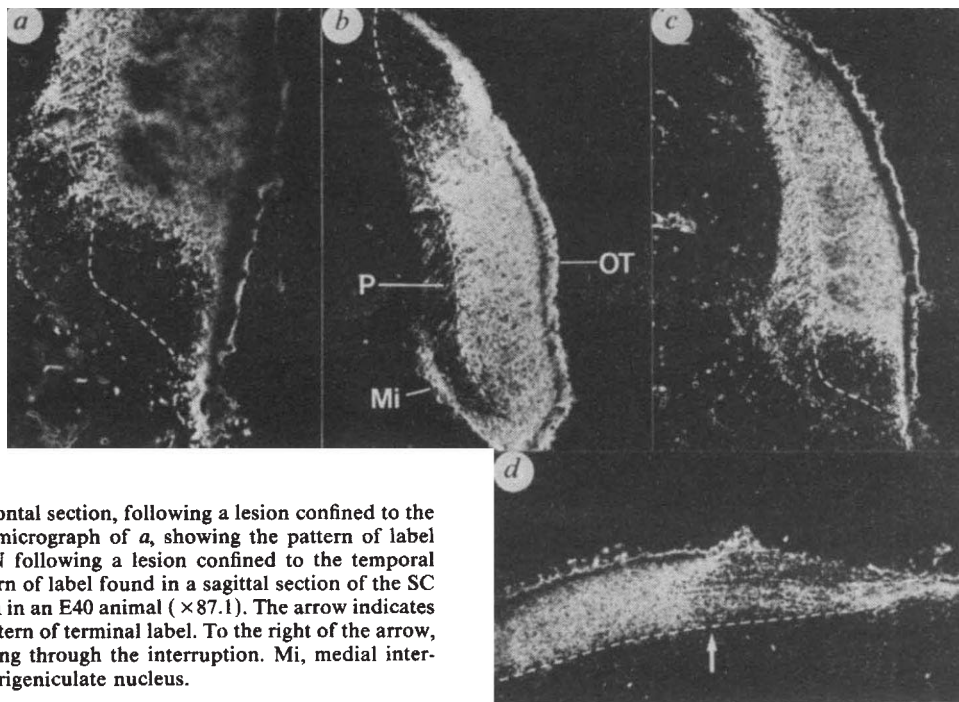
The topographical pattern found in the contralateral projection of the prenatal and neonatal animals is shown schematically in Fig. 2, together with a schematic representation of this projection in a pigmented adult.

Labelled terminals in the ipsilateral dorsal LGN and SC were very sparse in prenatal and albino animals, therefore it was not possible to determine the precision of this projection in these animals. However, in the neonatal pigmented animals, retinal lesions produced clear interruptions in the pattern of terminal label in the ipsilateral dorsal LGN, and the border between labelled and unlabelled regions was as sharp as that in the contralateral projection. Although this projection was topographically ordered, it was not possible to determine the nature of the topography.

I have demonstrated that the polarity of the contralateral projection in young ferrets is appropriate for the adult map.

* Permanent address: Department of Human Anatomy, University of Oxford, South Parks Road, Oxford OX1 3QX, UK.

Fig. 1 Dark-field photomicrographs of horizontal (*a-c*) and sagittal (*d*) sections of the dorsal LGN and SC from young ferrets after localized retinal lesions and anterograde labelling with HRP (50% in 0.9% saline); the tracer was injected intraocularly. In all cases the broken line marks the extent of the nuclei, determined using cytoarchitectonic criteria under bright-field illumination. The precision of the border between labelled and unlabelled regions is demonstrated in *a* (the sections shows the caudal region of the dorsal LGN from a D0 animal following a lesion of the contralateral temporal retina). HRP was injected after the lesion was made ($\times 134$). *b*, An example of the pattern of label found in the contralateral dorsal LGN of an E40 animal, shown in horizontal section, following a lesion confined to the nasal retina ($\times 87.1$). *c*, A low-power micrograph of *a*, showing the pattern of label found in the contralateral dorsal LGN following a lesion confined to the temporal retina ($\times 87.1$). *d*, Example of the pattern of label found in a sagittal section of the SC contralateral to a temporal retinal lesion in an E40 animal ($\times 87.1$). The arrow indicates the border of the interruption in the pattern of terminal label. To the right of the arrow, labelled optic fibres can be seen running through the interruption. Mi, medial interlaminar nucleus; OT, optic tract; P, perigeniculate nucleus.



However, two questions remain unanswered. First, although it has been shown that the temporal retina projects to the caudal part of the dorsal LGN, and nasal retina to the rostral part of the nucleus, it has not been possible to determine whether retinal regions between the two project in appropriate topographical order to intermediate parts of the nucleus. In spite of this, there

is no reason to assume that this region is different from either the rostral or caudal aspects. Second, in the adult there is a sharp division between ipsilaterally and contralaterally projecting ganglion cells which terminate in the LGN. In the neonate a small population of contralaterally projecting cells are found in the temporal retina⁸; it is unclear where these cells terminate in the dorsal LGN. Unfortunately, the temporal retina is relatively small in the ferret, and lesions in this region probably also encroached on the nasal retina. Hence, it was impossible to determine whether this aberrant projection terminated in topographical sequence in the caudal region of the nucleus.

The results presented here demonstrate that, during development, retinotopic order is established before ocular separation occurs in subcortical visual structures. It seems that retinotopic order is independent of any single set of landmarks in the dorsal LGN. If such landmarks were significant, two separate dynamic sets would be required, for which there is no evidence. While segregation is occurring, with one projection retracting to a smaller territory than the other, the two maps must slide in relation to one another, bringing the two projections into register in their separate territories. This process may be dependent on, or coincident with, retinal ganglion cell death. In the rodent, cell death in the population of ipsilaterally projecting cells occurs over a similar time period, and at a similar rate, to the retraction of the terminal field of ipsilaterally projecting cells in dorsal LGN². Unfortunately, no detailed study of the time course of cell death has been undertaken in the ferret, although preliminary studies indicate that it is a postnatal event in this species⁸.

I thank R. W. Guillery for his encouragement and support during this study, Ben Reese for commenting on the manuscript, and Steve Price for help with photography. This research was supported by grant PHS-NS-14283.

Received 24 September; accepted 30 November 1984.

1. Polyak, S. *The Vertebrate Visual System* (University of Chicago Press, 1957).
2. Jeffrey, G. *Dev Brain Res.* 13, 81-96 (1984).
3. Shatz, C. J. *J. Neurosci.* 3, 482-499 (1983).
4. Rakic, P. *Science* 214, 928-931 (1981).
5. Linden, D. C. et al. *J. comp. Neurol.* 203, 189-211 (1981).
6. Cucchiari, J. & Guillery, R. W. *Proc. R. Soc.* (in the press).
7. Messulam, M. M. *J. Histochem. Cytochem.* 26, 106-117 (1978).
8. Cucchiari, J. thesis, Univ. Chicago (1984).

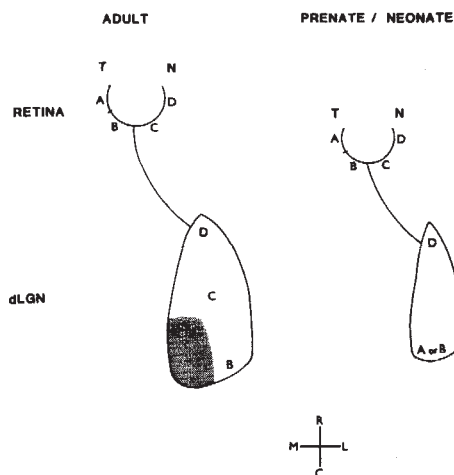


Fig. 2 Schematic representations of the contralateral retinal projection in the adult animal and in the late prenatal/early neonatal stage. In the adult retina regions B, C and D project in an orderly manner to the contralateral dorsal LGN (dLGN), with retinal region D projecting to the caudal part of the nucleus. There is a sharp division between ipsilaterally and contralaterally projecting ganglion cells between A and B. The shaded region in the adult dLGN represents the part of the nucleus occupied by terminals from the ipsilateral eye. The contralateral projection in the late prenatal/early neonatal animal covers the whole nucleus. Here, the nasal retina (D) is shown projecting to the rostral part of the nucleus as it does in the adult, with the temporal retina (A) projecting to the caudal part of the nucleus. Although this demonstrates that the unretracted projection is retinotopically ordered, it is unclear whether the interruption caused by the temporal retinal lesion was the result of damage to an aberrant crossed projection from A or damage to B. T, temporal; N, nasal; R, rostral; C, caudal; M, medial; L, lateral.

Functional innervation of cultured hippocampal neurones by cholinergic afferents from co-cultured septal explants

B. H. Gahwiler* & D. A. Brown†

* Preclinical Research, Sandoz Ltd, CH 4002 Basel, Switzerland and

† MRC Neuropharmacology Research Group, Department of Pharmacology, School of Pharmacy, University of London, 29/39 Brunswick Square, London WC1N 1AX, UK

The rat hippocampus receives a strong cholinergic innervation from the medial septum¹⁻³; information about the development and function of this pathway could help to elucidate the mechanisms of memory functions⁴⁻⁶. Previous electrophysiological studies have shown that septal stimulation *in vivo* facilitates commissural^{7,8} and perforant path⁹ inputs and that stimulation of intrahippocampal cholinergic fibres *in vitro* produces a slow depolarization of rat hippocampal CA₃ pyramidal neurones and increases their excitability^{10,11}. We describe here a different approach to the investigation of this system, by co-culturing slices of young rat hippocampus and septum¹², then recording the effects of septal nucleus stimulation on single voltage-clamped hippocampal CA₃ pyramidal neurones. Under these conditions acetylcholinesterase-staining (presumed cholinergic) fibres grow out from the septum into the hippocampus. Single septal stimuli produce a short-latency non-cholinergic fast excitatory postsynaptic current, whereas trains of stimuli produce a slow inward current augmented by neostigmine and suppressed by atropine; hence this has a cholinergic origin. Our experiments provide both the first demonstration that functional synapses can be established between explanted cholinergic and cholinceptive neuronal systems from the mammalian brain in organotypic culture and the first description of cholinergic slow excitatory postsynaptic currents in the mammalian central nervous system.

Slices of hippocampus and septum were obtained from 7-day-old rats and cultured side-by-side on glass coverslips^{13,14}. After 3-5 weeks the septal cultures contained groups of neurones

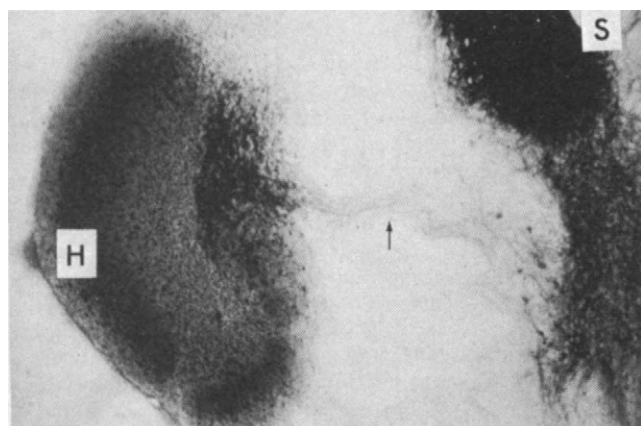


Fig. 1 Acetylcholinesterase-staining of a co-cultured septal (S) and hippocampal (H) slice preparation in which e.p.s.cs had been detected previously in a CA₃ pyramidal cell following a brief train of septal stimuli (see Fig. 2). Acetylcholinesterase-positive fibres (arrowed) originating in septal neurones cross the gap between the two explants and invade the dentate gyrus and the CA₃ area of the hippocampal formation.

Methods. The cultures were fixed in 4% paraformaldehyde in 0.15 M phosphate buffer and stained for acetylcholinesterase²⁵. The histochemistry used acetylthiocholine as the substrate and ethopropazine as the inhibitor of non-acetylcholinesterases.

showing a positive reaction to a cholinesterase stain and a proportion of the cultures had a strong outgrowth of cholinesterase-staining fibres from the septum into the hippocampus (Fig. 1). These fibres ramified widely throughout the hippocampus but were concentrated most heavily in the dentate gyrus and CA₃ areas.

For electrophysiological study, the co-cultures were transferred to a temperature-controlled (35 °C) perfusion chamber where single CA₃ hippocampal pyramidal cells were impaled with a microelectrode under visual control¹³ and then voltage-clamped¹⁵. Septal stimulation produced two different excitatory post-synaptic currents (e.p.s.cs): (1) a short-latency (<10 ms) fast e.p.s.c. after single shocks; (2) a delayed slow e.p.s.c. after a train of shocks (Fig. 2, left and right columns respectively).

The fast e.p.s.c. was recorded in most (20 of 24) of the co-cultures as a transient inward current of several nA amplitude

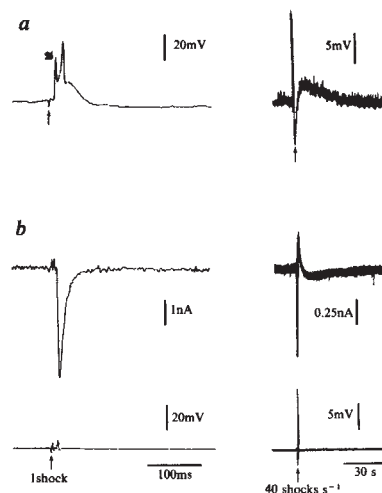


Fig. 2 Fast and slow excitatory postsynaptic potentials and currents (e.p.s.ps and e.p.s.cs) recorded from a CA₃ neurone in an organotypically-cultured hippocampal slice following a stimulation of the adjacent, co-cultured medial septum. Fast responses (left column) were evoked by a single septal stimulus; slow responses (right column) followed a 1-s train of stimuli at 40 Hz. Note the different amplifications and time-bases. *a*, Potential changes in the unclamped state (resting potential -42 mV, depolarization upwards); *b*, the corresponding membrane currents when the neurone was voltage-clamped at its resting potential (inward current downwards). The fast e.p.s.p. (arrowed) initiated by a single septal stimulus induced a brief 'burst' response²⁶ in the CA₃ neurone; under voltage clamp the burst is suppressed, leaving a monotonic fast e.p.s.c., indicating that the burst resulted from the activation of endogenous voltage-dependent currents by the initial synaptic depolarizations. The time-course of the slow e.p.s.p. induced by repetitive stimulations reflected faithfully the underlying slow e.p.s.c. (In this cell, the slow e.p.s.p./slow e.p.s.c. is preceded by an i.p.s.p./i.p.s.c., probably due to activation of intrahippocampal inhibitory interneurons.)

Methods. Co-cultures on a microscope slide were mounted on the stage of an inverted microscope and superfused at 1 ml min^{-1} with Hanks' balanced saline solution at 35 °C containing: 142 mM Na^+ ; 5.8 mM K^+ ; 154 mM Cl^- ; 2.8 mM Ca^{2+} ; 4.0 mM Mg^{2+} ; 4.2 mM HCO_3^- ; 0.4 mM SO_4^{2-} ; 5.6 mM D-glucose . The Mg^{2+} concentration was raised over the normal concentration (0.9 mM) to reduce spontaneous activity; in other tests, 8 mM Mg^{2+} with 4 mM Ca^{2+} was sometimes used. Neurones were impaled with a microelectrode filled with 3 M KCl (tip resistance $20\text{--}30 \text{ M}\Omega$) and voltage-clamped via a high frequency sample-and-hold amplifier (Axoclamp-2, Axon Instruments Inc.) switching between current injection and voltage recording at $2\text{--}6 \text{ KHz}$. Our techniques for culturing and recording from hippocampal neurones and applying this type of voltage-clamp have been described elsewhere¹³⁻¹⁵. Currents were filtered at $3\text{--}10 \text{ KHz}$ (for fast transients) or $0.3\text{--}1 \text{ KHz}$ (for slow currents) and recorded on a Gould 2400 recorder. Voltages were corrected for those recorded after withdrawing the electrode. The septum was stimulated via a monopolar glass microelectrode filled with 1 M NaCl , using $0.1\text{--}0.4 \text{ ms}$ square pulses of $1\text{--}10 \mu\text{A}$.

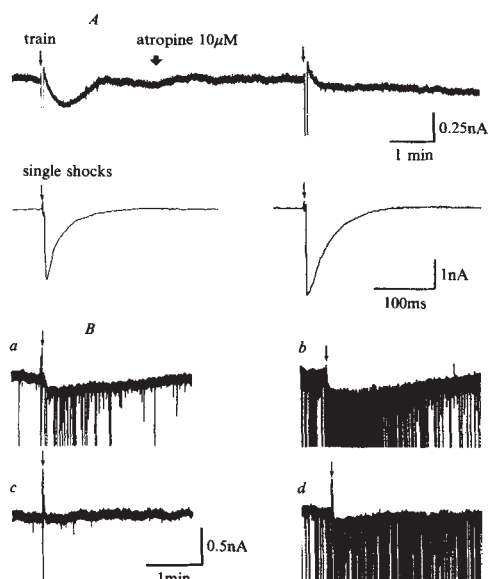


Fig. 3 The cholinergic nature of the slow e.p.s.c. as shown by the effects of atropine and neostigmine. **A**, Atropine (10 μ M) in the absence of neostigmine abolishes very rapidly the slow e.p.s.c. evoked by repetitive (40 Hz, 1 s) septal stimulation (upper traces: continuous time-trace) without inhibiting the fast e.p.s.c. evoked by a single septal stimulus (lower records). Holding potential -43 mV. **B**, Holding potential -31 mV; **a**, slow e.p.s.c.; **b**, slow e.p.s.c. enhanced after adding 1 μ M neostigmine; **c**, further addition of atropine (0.1 μ M) abolished the slow e.p.s.c.; **d**, partial recovery after removing the atropine. The rapid downward deflexions are spontaneous fast inward synaptic currents (see Fig. 2).

which lasted up to 100 ms and reversed in direction between -20 and -10 mV. It did not seem to be cholinergic in origin as it was not reduced by 10 μ M atropine or 100 μ M D-tubocurarine; instead, it was reduced by 1 mM kynurenic acid¹⁶ and so might be mediated by an amino acid. The adult septo-hippocampal pathway seems to contain many non-cholinergic fibres whose chemical identity has not yet been established^{17,18}.

The slow e.p.s.c. was detected in about half (10 of 24) of the cultures as a delayed inward current following a brief train of septal stimuli with a peak amplitude of a few hundred pA and lasting up to 3 min. In unclamped cells this slow inward current generated a slow depolarization similar to that described previously in adult hippocampal slices following stratum oriens stimulation^{10,11} (Fig. 2a). The slow e.p.s.c. seemed to be cholinergic because: (1) it was blocked readily by 0.1–10 μ M atropine (Fig. 3); (2) it was increased and prolonged by adding the anticholinesterase drug neostigmine (1 μ M; Fig. 3); (3) the effect of repetitive septal stimulation was imitated by cholinergic agonists (see Fig. 5). The slow e.p.s.c. was often accompanied by an increased frequency of spontaneous fast synaptic currents, especially in neostigmine solution, and was eliminated by atropine. In some cells, septal stimulation increased spontaneous e.p.s.c. frequency in the absence of an overt slow e.p.s.c. As these spontaneous currents were eliminated by atropine, they may result from an indirect cholinergic excitatory action on other hippocampal neurones connected synaptically to the impaled cell.

Membrane conductance changes during the slow e.p.s.c. were assessed by applying square voltage jumps before and immediately after septal stimulation in the presence of neostigmine (Fig. 4). The size of the current excursions was reduced appreciably during the slow e.p.s.c., indicating a fall in membrane input conductance, with a striking reduction in the magnitude of the time-dependent current relaxations observed during the voltage-jumps.

Two time- and voltage-dependent K^+ currents may contribute to these relaxations: a Ca^{2+} -independent K^+ -current, I_M ^{19,20}

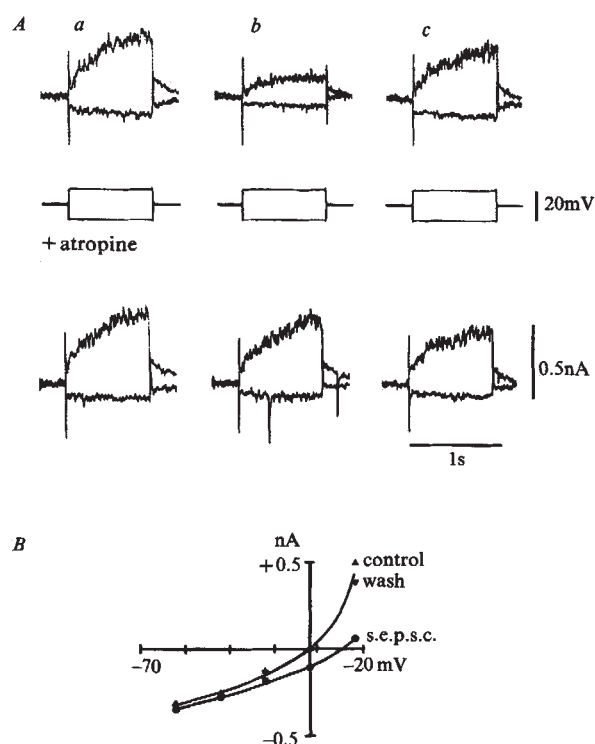
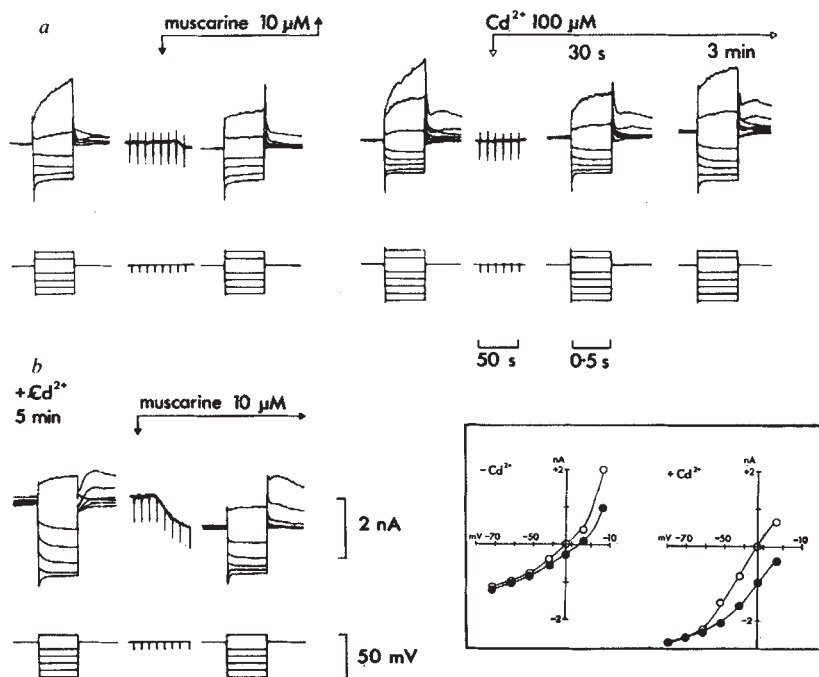


Fig. 4 Effect of septal stimulation on cell input conductance. **A**, A hippocampal CA₃ neurone was voltage-clamped at about -32 mV and **a**, subjected to 1-s voltage jumps of ± 10 mV; **b**, immediately after septal stimulation (during a slow e.p.s.c.) the currents evoked by the voltage jumps were reduced transiently in amplitude indicating a fall in input conductance; **c**, recovery 5 min later. Addition of 0.1 μ M atropine (lower records) inhibited the slow e.p.s.c. and prevented the effect of septal stimulation on input conductance. The culture was superfused with a solution containing 8 mM Mg^{2+} , 4 mM Ca^{2+} and 1 μ M neostigmine. The noise on the current records (filtered at 300 Hz) is due partly to spontaneous (non-cholinergic) synaptic activity. **B**, The current-voltage relationship for this cell measured before (Δ), during ($\bullet$) and after (∇) the slow e.p.s.c. Currents were measured at the end of a series of 1-s commands of varying amplitude, as displacement from the holding current at -32 mV. Each point is the mean of three measurements after three separate bursts of septal stimulation.

and a Ca^{2+} -activated K^+ current, I_C ¹⁵. (There may be contamination also from a Ca^{2+} -activated Cl^- current²¹.) These currents could not be separated adequately by pharmacological means while preserving the septal stimulation protocol. Tests using the cholinomimetic drug muscarine²⁰, however, suggested that inhibition of the Ca^{2+} -independent current, I_M , was the major cause of both the fall in input conductance and the inward synaptic current (Fig. 5). Thus, the action of muscarine on both steady current level and voltage-induced current relaxations persisted when inward Ca^{2+} -currents were blocked with 100 μ M Cd^{2+} . These effects of muscarine were imitated by acetylcholine (1 μ M) and Ba^{2+} (1 mM) and were blocked in the presence of the Ba^{2+} ions, in agreement with previous observations on the properties of I_M ^{19,20}. Further support for the view that I_M inhibition provides the driving force for the depolarizing inward current is provided by the current-voltage curves in Figs 4 and 5, as these converge at hyperpolarized command potentials where I_M becomes deactivated. Likewise, synaptically-driven slow e.p.s.c.s could only be detected at membrane holding potentials positive to about -55 mV and increased in amplitude with membrane depolarization, *pari passu* the activation of I_M .

In adult rat hippocampal slices, stratum oriens stimulation seems to inhibit a Ca^{2+} -activated K^+ -conductance responsible for spike adaptation and post-spike hyperpolarization^{10,22}, an effect imitated by noradrenaline²³, 10 μ M of which had a similar effect on our cultured cells. Unlike septal stimulation or

Fig. 5 Effects of muscarine and Cd^{2+} on clamp-currents in a CA_3 neurone, superfused with a solution containing $1 \mu\text{M}$ tetrodotoxin (TTX) to suppress indirect (trans-synaptic) effects of muscarine. The records show superimposed current responses to several series of 0.5-s square voltage commands similar to those illustrated in Fig. 4. **a**, $10 \mu\text{M}$ DL-muscarine produced an inward current and reduced input conductance similar to septal stimulation (see Fig. 4). After washing out the muscarine for 5 min (two current responses to $+20\text{ mV}$ commands are shown to indicate the range of variation), $100 \mu\text{M}$ CdCl_2 was added to the superfusion fluid for the remainder of the experiment. This concentration of Cd^{2+} eliminated inward Ca^{2+} -currents in other cells bathed in TTX-solution and impaled with a Cs^+ -filled microelectrode²⁷. Cd^{2+} reduced initially the amplitude of the larger outward current relaxations but not the smaller outward relaxations, or the inward relaxations produced by hyperpolarizing commands. After 3-min exposure to Cd^{2+} , a sustained outward current developed accompanied by an increased conductance (regularly observed with Cd^{2+}). Relaxations in Cd^{2+} solution accelerated with increasing hyperpolarization and inverted at -80 mV , near to our estimate of E_K in these cells when bathed in 6 mM K^+ (ref. 27). (The 'rebound' outward currents after large hyperpolarizing commands could be blocked by $100 \mu\text{M}$ 4-aminopyridine and so may be inactivating K^+ -currents²⁸). **b**, After 5-min perfusion with Cd^{2+} , $10 \mu\text{M}$ muscarine was reapplied and again produced an inward current with a fall in input conductance. The plots (inset) show current-voltage curves measured from the current responses to the voltage-steps recorded before ($\circ$) and during ($\bullet$) the applications of muscarine in the absence and presence of Cd^{2+} .



cholinomimetic drug application, however, noradrenaline did not inhibit the time-dependent current relaxations of Figs 4 and 5 and did not induce an inward current. Hence, even if contributing to excitability changes following septal stimulation, inhibition of this Ca^{2+} -activated current cannot account for the slow synaptic current itself, a view substantiated by the fact that no inward current was generated by Cd^{2+} ions (see Fig. 5).

We conclude that the cholinergic slow e.p.s.c. recorded in these cultured hippocampal cells following septal stimulation is analogous to that described previously in some sympathetic ganglia²⁴ and results from the ability of the released acetylcholine to inhibit the voltage-dependent outward current I_M . In more general terms, our success in establishing functional synapses between septal and hippocampal cells in culture suggests that this technique offers a useful approach to studying transmission between other remote brain loci.

Supported in part by the MRC. We thank Frl. E. Hoffmann and L. Rietschin for their help and Mrs. Jenny Hollington for processing the manuscript.

Received 13 June; accepted 15 November 1984.

- Lewis, P. R. & Shute, C. C. D. *Brain* **90**, 521-542 (1967).
- Mesulam, M.-M., Mufson, E. J., Wainer, B. H. & Levey, A. I. *Neuroscience* **19**, 1185-1201 (1983).
- Milner, T. A., Loy, R. & Amaral, D. G. *Dev. Brain Res.* **8**, 343-371 (1983).
- O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Clarendon, Oxford, 1978).
- Olton, D. S., Becker, J. T. & Handelmann, G. E. *Behav. Brain Sci.* **2**, 313-365 (1979).
- Dunnett, S. B. *et al. Brain Res.* **251**, 335-348 (1982).
- Krnjevic, K. & Ropert, N. *Can. J. Physiol. Pharmacol.* **59**, 911-914 (1981).
- Krnjevic, K. & Ropert, N. *Neuroscience* **7**, 2165-2183 (1982).
- Fantie, B. D. & Goddard, G. V. *Brain Res.* **252**, 227-237 (1982).
- Cole, A. E. & Nicoll, R. A. *Science* **221**, 1299-1301 (1983).
- Cole, A. E. & Nicoll, R. A. *J. Physiol., Lond.* **352**, 173-188 (1984).
- Gähwiler, B. H. & Hefti, F. *Neuroscience* **13**, 681-689 (1984).
- Gähwiler, B. H. *J. Neurosci. Meth.* **4**, 329-342 (1981).
- Gähwiler, B. H. *Experientia* **40**, 235-243 (1984).
- Brown, D. A. & Griffith, W. H. *J. Physiol., Lond.* **337**, 287-301 (1983).
- Perkins, M. N. & Stone, T. W. *Brain Res.* **247**, 184-187 (1982).
- Mesulam, M., Van Hoesen, G. & Rosene, D. G. *Soc. Neurosci. Abstr.* **3**, 202 (1977).
- Baisden, R. H., Woodruff, M. L. & Hoover, D. A. *Brain Res.* **290**, 146-151 (1984).
- Brown, D. A. & Adams, P. R. *Nature* **283**, 673-676 (1980).
- Halliwel, J. V. & Adams, P. R. *Brain Res.* **250**, 71-92 (1982).
- Owen, D. G., Segal, M. & Barker, J. L. *Nature* **311**, 567-570.
- Lancaster, B. & Adams, P. R. *Neurosci. Abstr.* **10**, 872 (1984).
- Madison, D. V. & Nicoll, R. A. *Nature* **299**, 636-638 (1982).
- Adams, P. R. & Brown, D. A. *J. Physiol., Lond.* **332**, 329-342 (1982).
- Geneser-Jensen, F. A. & Blackstad, T. W. *Zellforsch.* **114**, 460-481 (1972).
- Spencer, W. A. & Kandel, E. R. *J. Neurophysiol.* **24**, 272-285 (1961).
- Gähwiler, B. H. & Brown, D. A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Gustafsson, B., Galvan, M., Grafe, P. & Wigstrom, H. *Nature* **299**, 252-254 (1982).

Light-induced reduction of cytoplasmic free calcium in retinal rod outer segment

King-Wai Yau & Kei Nakatani

Department of Physiology and Biophysics,
University of Texas Medical Branch, Galveston, Texas 77550, USA

The response of retinal rod photoreceptors to light consists of a membrane hyperpolarization resulting from the decrease of a light-sensitive conductance in the outer segment¹. According to the calcium hypothesis, this conductance is blocked by a rise in intracellular free Ca triggered by light², a notion supported by the findings that an induced rise in internal Ca leads to blockage of the light-sensitive conductance³⁻⁹ and that light triggers a net Ca efflux from the outer segment via a Na-Ca exchanger, suggesting a rise in internal free Ca in the light¹⁰⁻¹³. We have now measured both Ca influx and efflux through the outer segment plasma membrane and find that, contrary to the calcium hypothesis, light seems to decrease rather than increase the free Ca concentration in the rod outer segment. This result implies that Ca does not mediate visual excitation but it probably has a role in light adaptation.

Figure 1A shows the responses of a toad (*Bufo marinus*) rod to flashes (traces a,c) and steps (traces b,d) of light. All responses were saturated, but the flash in c and the light step in d were 60-110 times brighter than the corresponding ones in a and b. A striking feature is that all the responses exhibited, at their initial plateau, a small secondary climb whose amplitude and time course were independent of light intensity and duration. This climb or 'transient' apparently reflected the decline of an electrogenic Na -dependent Ca efflux at the outer segment because it disappeared (Fig. 1B) when all external Na around the outer segment was replaced with Li , a cation incapable of driving the Ca efflux¹⁴. From 28 experiments, the Na-Ca exchange current (an inward current) was determined to have an initial peak of $1.5 \pm 0.3\text{ pA}$ (mean $\pm$ s.d.) followed by an

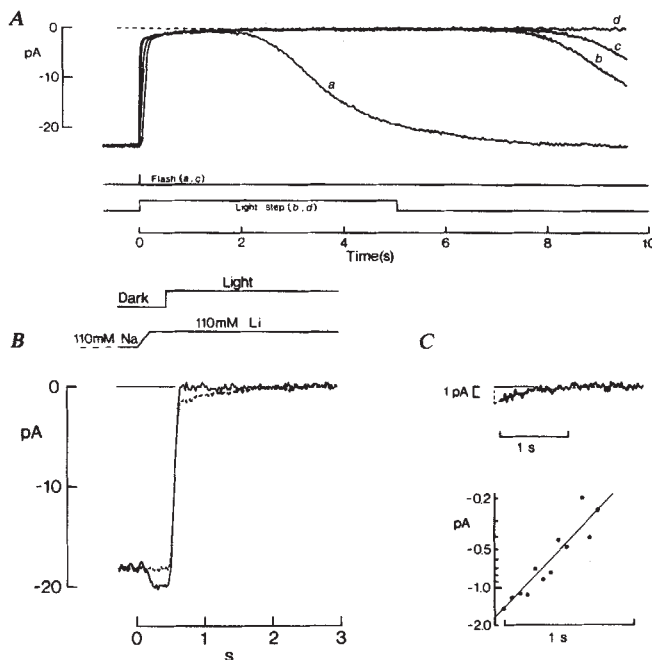


Fig. 1 **A**, Flash and step responses of a toad rod to light. Under infrared light a piece of retina from a dark-adapted toad was finely chopped in Ringer's solution, yielding some isolated rods with intact outer and inner segments³⁰. Membrane current from one of these rods was recorded by sucking its outer segment into a close-fitting glass pipette containing Ringer's solution and connected to a current/voltage converter. Normal Ringer's solution contained (in mM): 110 Na, 2.5 K, 1.6 Mg, 1.0 Ca, 117.7 Cl, 5 TMA (tetramethylammonium)-HEPES, 5 dextrose, pH 7.6. The inner segment was perfused also with Ringer's solution, but this contained in addition 1 mM CsCl to remove the initial hyperpolarizing transient in the voltage response to light^{31,32}. We found that in very bright light this transient causes a capacity current which partially obscured the Na-Ca exchange current at early times of the response plateau; hints of this capacity current can be detected in the experiments of **B** (see below) and Fig. 2, in which no CsCl was present. Outward membrane current across the outer segment is plotted upwards. Responses *a* and *c* were elicited by flashes (7.6-ms duration), causing $\sim 4.8 \times 10^2$ and 5.4×10^4 photoisomerizations (Rh^*). Responses *b* and *d* were elicited by light steps (~ 5 s duration), producing $\sim 8.1 \times 10^3$ and $4.7 \times 10^5 Rh^* s^{-1}$. An effective collecting area of $15 \mu m^2$ for the rod outer segment was assumed throughout with unpolarized light (see ref. 17). Each response is an average of three sweeps. All recordings were low-pass-filtered at 100 Hz, 8 pole; all light stimuli were at 500 nm and unpolarized. Room temperature. **B**, Substitution of Li for Na around a rod outer segment removes the Na-Ca exchange current transient at the light response plateau. The recording was made with the inner segment of the rod within the suction pipette (containing normal Ringer's solution) and its outer segment perfused^{7,8,30}. Solid trace, response to a saturating light step in Li. Broken trace, control response (to the same light) in normal Ringer's. The time course of solution change, as indicated by junction currents⁷, was ~ 200 ms and is shown above the sweeps. The valve controlling the solution changes was operated pneumatically and with electronic timing so that the solution change could be reproduced exactly several times. The solid trace was obtained from the total recorded current (physiologically current plus junction current) measured in a dark/light-step sequence minus the junction current measured in a repeated run in continuous light. Each trace is the mean of three sweeps. The Li solution is identical to normal Ringer's in composition except for equimolar substitution of Li for Na. The saturating light step produced $\sim 2.6 \times 10^4 Rh^* s^{-1}$. The initial increase in inward dark current upon introduction of Li is due to the greater permeability for Li (as compared to Na) of the light-sensitive conductance^{8,9}. **C**, Upper, difference between control and Li responses. Broken curve, exponential decline with a time constant of 0.39 s; exponential extrapolation back to the time of 50% saturated light response gives a peak amplitude of 1.7 pA. Lower, semilogarithmic plot of the same trace (filtered at 25 Hz, 8 pole) at 60-ms intervals. The straight line drawn through the points gives a time constant of 0.39 s.

approximately exponential decline (Fig. 1C) with a time constant of 0.4 ± 0.1 s (mean $\pm$ s.d.) to a residual level below the limit of resolution (~ 0.1 pA) of our recording system. Intense flashes superposed on a just-saturating steady response failed to produce any incremental jumps in exchange current above the residual level (data not shown), suggesting that the transient Ca efflux at the onset of illumination was associated with the light-sensitive conductance shutting down rather than with light *per se*.

As the light-sensitive conductance is believed to be permeable to Ca^{2+} , the transient Ca efflux in the light may simply reflect a high Ca influx in the dark, in which case if the dark Ca influx were stopped by removing external Ca, the extrusion would also quickly subside and a subsequent light response should lack the transient. The experiment of Fig. 2A confirmed this and indicated that the ability of light to shut down the conductance need not be associated with any Ca efflux. Figure 2B shows the reverse experiment, in which the dark current, and hence the Ca influx, in Ringer's was increased by the presence of the phosphodiesterase inhibitor, isobutylmethylxanthine (IBMX)^{7,15,16}. As expected, the light response showed a progressively larger transient at its plateau (solid traces). This large transient again disappeared when Li replaced external Na (dashed trace), indicating that it represented heightened Na-dependent Ca efflux. The S-shaped instead of purely exponential decline exhibited by the larger transients suggested that as the internal free Ca concentration increased substantially, the rate of Ca extrusion gradually approached saturation. The maximal exchange current observed in Ringer's with large Ca loads was ~ 30 pA.

Figure 3 shows an experiment to estimate the dark Ca influx in normal conditions, in which we trapped the dark Ca influx within the outer segment by removing external sodium and estimated the influx from the subsequently reactivated efflux¹⁴ (see Fig. 3 legend). We estimated that ~ 10 – 15% (5 experiments) of the total normal dark current is carried by Ca, or 2.0–3.0 pA for a dark current of 20 pA. At steady state such an influx would be extruded completely via the Na-Ca exchanger, which, by virtue of its 3:1 stoichiometry¹⁴, would produce an inward exchange current of 1.0–1.5 pA, a value in rather good agreement with the peak of the transient (~ 1.5 pA) at the response plateau described earlier. An interesting corollary is that the influx of Na via the exchanger in the dark should be ~ 4.5 pA ($= 3 \times 1.5$ pA), which is as much as 25% of the Na influx through the light-sensitive conductance. For a given state of the light-sensitive conductance, we found that the Ca influx remained roughly the same whether the Na in Ringer's was replaced by a permeant (Li or guanidinium) or an impermeant (choline) cation. Thus, at least at physiological concentrations, the inward movement of Ca through the conductance seems to be largely independent of the other external cations.

A simple picture that emerges is that in the dark Ca steadily enters the rod outer segment at a rather high rate ($\sim 10^7$ ions s^{-1}), to be balanced by an equal efflux via the Na-Ca exchanger. Bright light completely shuts down the influx; the efflux continues and perhaps increases because of the hyperpolarization and the exchange's electrogenicity, but very rapidly declines when the internal free Ca concentration falls. The approximately exponential decline of the exchange current supports the idea that the Ca extrusion rate is more or less linearly related to internal free Ca concentration and that the free Ca extruded is not readily replenished by release of bound Ca. This apparent lack of rapid equilibrium between free and bound Ca means that any Ca influx trapped in the outer segment should remain free and accelerate the exchange, as observed (Fig. 3).

One can predict the net Ca efflux induced by a single photoisomerization (Rh^*). A photon inhibits the dark current by ~ 1 pA and with an integration time of ~ 1 s¹⁷, giving a total blockage of charge of ~ 1 pC; as 10–15% of this charge is thought to be carried by Ca ions, the photon effectively blocks the entry, or causes a net 'efflux', of 3×10^5 Ca ions. Perhaps more accurately, as a photon completely suppresses the dark current for

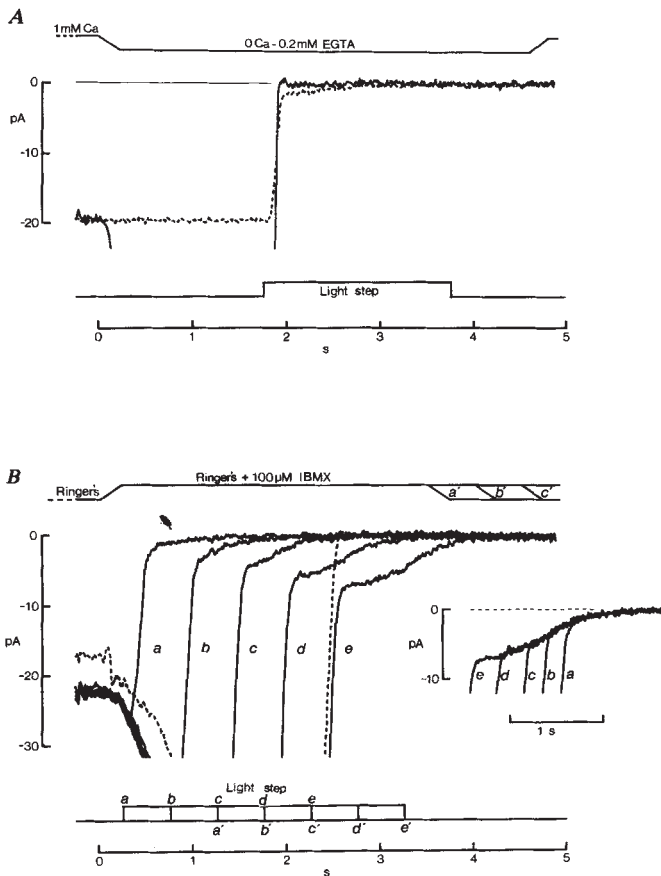


Fig. 2 **A**, Brief exposure of a rod outer segment to a solution containing no Ca removes the transient Na-Ca exchange current at the response plateau. The rod inner segment was within the pipette and the outer segment was perfused. Solid trace (single trial), the inward dark current rapidly increased to >250 pA on removing the external Ca around the outer segment and the subsequent response to a bright light step lacked the exchange current transient. The Na load due to the large dark current before illumination was estimated to be ~ 247 pC; at most this would increase the internal Na concentration by <3 mM. The 0-Ca solution was identical to normal Ringer's in composition except that Ca was omitted and 0.2 mM TMA-EGTA was added. Broken trace (average of 2 trials), the control response in normal Ringer's. The light step produced $\sim 1.6 \times 10^4$ $\text{Rh}^* \text{s}^{-1}$ in Ringer's and 4.5×10^5 $\text{Rh}^* \text{s}^{-1}$ in the 0-Ca solution. **B**, Brief exposure of rod outer segment to Ringer's containing $100 \mu\text{M}$ IBMX produces a larger dark current and a much larger transient exchange current than the control on illumination (different cell from **A**). Solid traces (single trial for each), experiment in normal Ringer's with IBMX, that is, Na-Ringer's $\rightarrow$ Na-Ringer's + IBMX. Broken trace (single trial), subsequent experiment on the same cell in Li-Ringer's with $100 \mu\text{M}$ IBMX, that is, Na-Ringer's $\rightarrow$ Li-Ringer's + IBMX. Inset, exchange current transients shifted for alignment and superposed. In the light monitor trace (bottom), *a-e* indicate turning on, while *a'-e'* indicate turning off, of respective light steps. The light produced $\sim 6.1 \times 10^4$ $\text{Rh}^* \text{s}^{-1}$ in each case.

~ 1 s over 3% of the outer segment length¹⁸, the resulting net Ca efflux would be 3% of the area under the exchange current in Fig. 1C, integrated over 1 s; this gives $\sim 10^5$ Ca ions per Rh^* . In either case, the predicted efflux is more than enough to account for measurements made by others from toad retina using external Ca-sensitive electrodes ($\sim 1-2 \times 10^4$ Ca ions per Rh^*)^{11,12}. With brighter flashes the total net Ca efflux should initially increase linearly with light level but at progressively higher intensities should rise more slowly, as the internal free Ca becomes depleted and the Ca extrusion declines. During the recovery of the dark current after a flash, the Ca movement should transiently be a net influx, because internal free Ca has

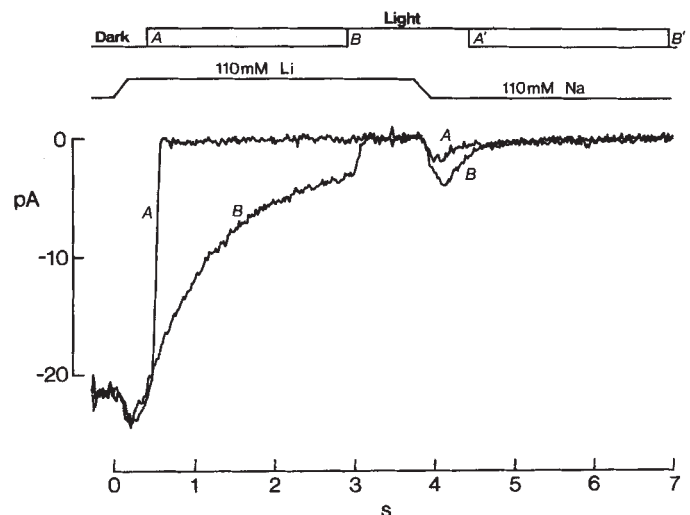


Fig. 3 Estimate of dark Ca influx obtained by trapping the Ca in the rod outer segment with a 0-Na, 110 mM Li solution, followed by reactivated extrusion in Na-Ringer's. The rod inner segment was in the pipette and the outer segment perfused. Except for equimolar substitution of Li for Na, the Li solution was otherwise identical to normal Ringer's. Two separate trapping measurements were made on the same cell, and this was controlled by light which shut off the light-sensitive conductance. The longer trapping period (trace B) led to a larger exchange current on readmission of external Na. The difference in total charge influx, ΔQ_{load} , between the two trapping periods (that is, the area between traces A and B during the time in the Li solution) was measured as 19.0 pC. The difference in subsequent charge transfer via the Na-Ca exchanger, ΔQ_{exch} (that is, the area between traces A and B on readmission of Na) was 0.89 pC. The incremental Ca load, ΔQ_{Ca} , is given by $\Delta Q_{\text{exch}} \times [2/(n-2)]$, where n is the stoichiometry of the exchange, previously measured to be 3 (ref. 14). The fractional Ca influx in the Li-Ringer's, f'_{Ca} , is then given by $f'_{\text{Ca}} = \Delta Q_{\text{Ca}}/\Delta Q_{\text{load}} = 0.094$. As the inward dark current is predominantly carried by Li ions, and since Li is 1.0-1.4 times as permeable as Na^{8,9}, the fractional Ca influx in normal Ringer's, f_{Ca} , would be 0.09-0.13. (The higher end of this estimated range is probably more accurate because the previous estimates for the apparent Li/Na selectivity ratio^{8,9} were not corrected for any exchange current component in the total dark current in Na-Ringer's; this would lead to a slight underestimate of the Li/Na ratio.) From six experiments, f_{Ca} was measured to be 0.09-0.15. In the light monitor trace (top) A and B indicate turning on, and A' and B' indicate turning off, of respective light steps. Steady light in this experiment produced 2.6×10^4 $\text{Rh}^* \text{s}^{-1}$. The incremental Ca loading method used here eliminates any error caused by the finite substitution period of Li for Na in the dark (during which some Ca entering the outer segment would still be extruded), and also any error in estimating the regular efflux in the light without any Ca load (which would have to be subtracted from the total efflux if either trace A or B alone were used for the above computations).

to be replenished by the resumed Ca influx before extrusion can be restored to dark level. Such an expected transient reversal in net Ca movement has indeed been observed¹⁹. In continuous bright light the net Ca efflux should, as demonstrated earlier, decline very rapidly from a peak of $\sim 10^7$ Ca ions s^{-1} to a basal level of either zero or at any rate $<6 \times 10^5$ Ca ions s^{-1} , as calculated from the upper limit of a 0.1 pA residual exchange current mentioned above. Extrusion may persist for an extended period of time if bound Ca starts to become unbound when internal free Ca reaches a critically low level. Assuming that the exchange current does decline exponentially to zero in bright light, the very minimal total Ca efflux would be $\sim 4 \times 10^5$ Ca ions; this should remove most of the free Ca within the outer segment if the free concentration in the dark is a few micromolar³. The maximal initial Ca efflux rate we found, $\sim 10^7$ Ca ions s^{-1} , is comparable to that reported by Schröder and Fain^{12,13} but while they apparently observed a high extrusion rate maintained

with time, we observed an immediate rapid decline in the rate in continuous light. The reason for this discrepancy is unknown.

Our results do not support the long-held calcium hypothesis that the free Ca concentration in the rod outer segment rises on illumination. On the contrary, as the Ca efflux declines immediately after onset of illumination, the free Ca level should decrease in the outer segment. This assumes, of course, that light itself does not inhibit the Na-Ca exchanger, which seems valid because we found that the rate of extrusion of any Ca trapped in the outer segment was the same in the light as in the dark (data not shown).

Although internal Ca clearly inhibits the light-sensitive conductance³⁻⁹, we and others have found recently that this does not appear to be caused by Ca binding directly to and blocking the conductance^{9,14}. Therefore it seems that Ca does not mediate phototransduction at the beginning, nor at the end, of the reaction cascade, leaving cyclic GMP as the most likely candidate²⁰. The question remains of what role, if any, Ca has in

transduction. There is evidence that a high concentration of Ca in the rod outer segment tends to suppress the level of cyclic GMP²¹⁻²³, perhaps by modulating guanylate cyclase²⁴⁻²⁷ or phosphodiesterase²⁸. If cyclic GMP maintains the open state of the conductance, then a rise in Ca should tend to shut it down, as has been observed³⁻⁹. On the other hand, a reduction of free Ca in the light should oppose the light-activated decline in the level of cyclic GMP; this negative feedback would therefore constitute a form of light adaptation. Indeed, such a possibility has been raised previously by others²⁹.

We thank Drs Denis Baylor, Malcolm Brodwick, Arthur Brown, Geoff Gold, Lawrence Haynes, L. E. Moore, Brian Nunn and Andrew Sömlyo for discussions and comments and Dr Kenneth Nielsen for help in computer programming. The possibility that the early transient in the rod response plateau represents electrogenic Na-Ca exchange activity was raised initially by Professor Alan Hodgkin. This work was supported by the US National Eye Institute and the Retina Research Foundation, Houston, Texas.

Received 8 October; accepted 19 December 1984.

1. Toyada, J.-I., Nosaki, H. & Tomita, T. *Vision Res.* **9**, 453-463 (1969).
2. Hagins, W. A. *Rev. Biophys. Bioengng* **1**, 131-158 (1972).
3. Hagins, W. A. & Yoshikami, S. *Expl Eye Res.* **18**, 299-305 (1974).
4. Brown, J. E., Coles, J. A. & Pinto, L. H. *J. Physiol., Lond.* **269**, 707-722 (1977).
5. Bastian, B. L. & Fain, G. L. *J. Physiol., Lond.* **330**, 331-347 (1982).
6. MacLeish, P. R., Schwartz, E. A. & Tachibana, M. *J. Physiol., Lond.* **348**, 645-664 (1984).
7. Hodgkin, A. L., McNaughton, P. A., Nunn, B. J. & Yau, K.-W. *J. Physiol., Lond.* **350**, 649-680 (1984).
8. Yau, K.-W. & Nakatani, K. *Nature* **309**, 352-354 (1984).
9. Hodgkin, A. L., McNaughton, P. A. & Nunn, B. J. *J. Physiol., Lond.* (in the press).
10. Yoshikami, S., George, J. S. & Hagins, W. A. *Nature* **286**, 395-398 (1980).
11. Gold, G. H. & Korenbrot, J. I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5557-5561 (1980).
12. Schröder, W. H. & Fain, G. L. *Nature* **309**, 268-270 (1984).
13. Schröder, W. H. & Fain, G. L. *Biophys. J.* **45**, 341a (1984).
14. Yau, K.-W. & Nakatani, K. *Nature* **311**, 661-663 (1984).
15. Lipton, S. A., Rasmussen, H. & Dowling, J. E. *J. gen. Physiol.* **70**, 771-791 (1977).
16. Capovilla, M., Cervetto, L. & Torre, V. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6698-6702 (1982).

17. Baylor, D. A., Lamb, T. D. & Yau, K.-W. *J. Physiol., Lond.* **288**, 613-634 (1979).
18. Lamb, T. D., McNaughton, P. A. & Yau, K.-W. *J. Physiol., Lond.* **319**, 463-496 (1981).
19. Gold, G. H. *A. M. Soc. Neurosci.* **10**, 621 (1984).
20. Miller, W. H. (ed.) *Molecular Mechanisms of Photoreceptor Transduction* (Academic, New York, 1981).
21. Cohen, A. I., Hall, I. A. & Ferrendelli, J. A. *J. gen. Physiol.* **71**, 595-612 (1978).
22. Kilbride, P. J. *J. gen. Physiol.* **75**, 457-465 (1980).
23. Woodruff, M. L. & Fain, G. L. *J. gen. Physiol.* **80**, 537-555 (1982).
24. Krishnan, N., Fletcher, R. T., Chader, G. J. & Krishna, G. *Biochim. biophys. Acta* **523**, 506-515 (1978).
25. Troyer, E. W., Hall, I. A. & Ferrendelli, J. A. *J. Neurochem.* **31**, 825-833 (1978).
26. Fleischman, D. & Denisevich, M. *Biochemistry* **18**, 5060-5066 (1979).
27. Lolley, R. N. & Racz, E. *Vision Res.* **22**, 1481-1486 (1982).
28. Robinson, P. R., Kawamura, S., Abramson, B. & Bownds, M. D. *J. gen. Physiol.* **76**, 631-645 (1980).
29. Bownds, M. D. *Photochem. Photobiol.* **32**, 487-490 (1980).
30. Yau, K.-W., McNaughton, P. A. & Hodgkin, A. L. *Nature* **242**, 502-505 (1981).
31. Fain, G. L., Quandt, F. N., Bastian, B. L. & Gershenfeld, H. M. *Nature* **272**, 467-469 (1978).
32. Torre, V. *J. Physiol., Lond.* **333**, 315-341 (1982).

Effects on the photoresponse of calcium buffers and cyclic GMP incorporated into the cytoplasm of retinal rods

H. R. Matthews*, V. Torre† & T. D. Lamb*‡

* Physiological Laboratory, University of Cambridge, Downing Street, Cambridge CB2 3EG, UK

† Dipartimento di Fisica, Università di Genova, Genova, Italy

It is generally accepted that the light response in retinal rods involves a reduction of ionic permeability (predominantly to Na⁺) in the plasma membrane of the outer segment and that this is mediated by an internal messenger which diffuses between the disk and plasma membranes. There is controversy, however, over the identity of the diffusible substance; two alternative schemes have received widespread support (for review see refs 1, 2). According to the 'calcium hypothesis'^{3,4}, light stimulates the release into the cytoplasm of calcium, leading to the blockage of channels which are normally open in darkness, whereas based on the 'cyclic nucleotide hypothesis'^{5,6}, cyclic GMP causes the opening of channels in the dark, but is hydrolysed by a light-activated phosphodiesterase. We report here effects of introducing calcium buffers and cyclic GMP into the rod cytoplasm by means of a patch pipette⁷, which seem to be inconsistent with the calcium hypothesis.

The suction pipette method^{8,9} was used to record the photocurrent from the outer segment (or sometimes the inner segment) of a single rod isolated, without the use of enzymes, from the retina of the larval salamander *Ambystoma tigrinum*. In this way it was possible to select only those rods which functioned normally and gave large responses to light (suction pipette current ≥ 35 pA). A patch pipette, containing artificial intracel-

lular solution¹⁰ together with any substance to be introduced, was sealed against the part of the rod protruding from the suction pipette (usually the inner segment) and was used to obtain a whole-cell recording⁷ by rupturing the membrane patch.

Figure 1 shows the suction pipette current recorded from a rod outer segment (ROS) on rupturing the membrane in this way with a patch pipette on the inner segment containing no added calcium and 10 mM of the calcium chelator 1,2-bis(*o*-aminophenoxy)ethane-*N,N,N',N'*-tetraacetic acid (BAPTA; see ref. 11 and Fig. 1 legend). Figure 1a shows a continuous record while b shows superimposed responses to bright flashes on a faster time scale. The arrow in Fig. 1a indicates the time at which the whole-cell recording was obtained; beginning at this time the contents of the pipette and of the cytoplasm presumably began exchanging through the ruptured patch. Soon after this the mean 'dark current' began increasing and the responses to bright flashes remained saturated for progressively longer (see Fig. 1b), exhibiting an overshoot and damped oscillation, both of which became progressively more pronounced. Despite the presumed diffusion of the buffer into the outer segment, the bright flashes continued to suppress completely the inward dark current. In other experiments (not shown) the dark current was suppressed by continuous bright illumination for at least 15 min (the duration of the recording), but when the light was turned off the current re-appeared with a time course similar to the recovery following flashes in Fig. 1, and bright light could again rapidly suppress it. As shown in Fig. 1b, the initial rate of rise of the bright flash response remained unchanged for at least 2 min after introduction of the buffer; similar effects have been observed in 27 successful whole-cell experiments using 2 or 10 mM BAPTA or 4 mM EGTA with free Ca²⁺ in the pipette buffered to $\leq 10^{-8}$ M. In control experiments using only trace amounts (20 μ M) of BAPTA or EGTA, the suction pipette photocurrent showed little change on rupturing the patch, although a gradual reduction in amplitude and slowing of the response was usually observed over a period of 10-20 min.

‡ To whom correspondence should be addressed.

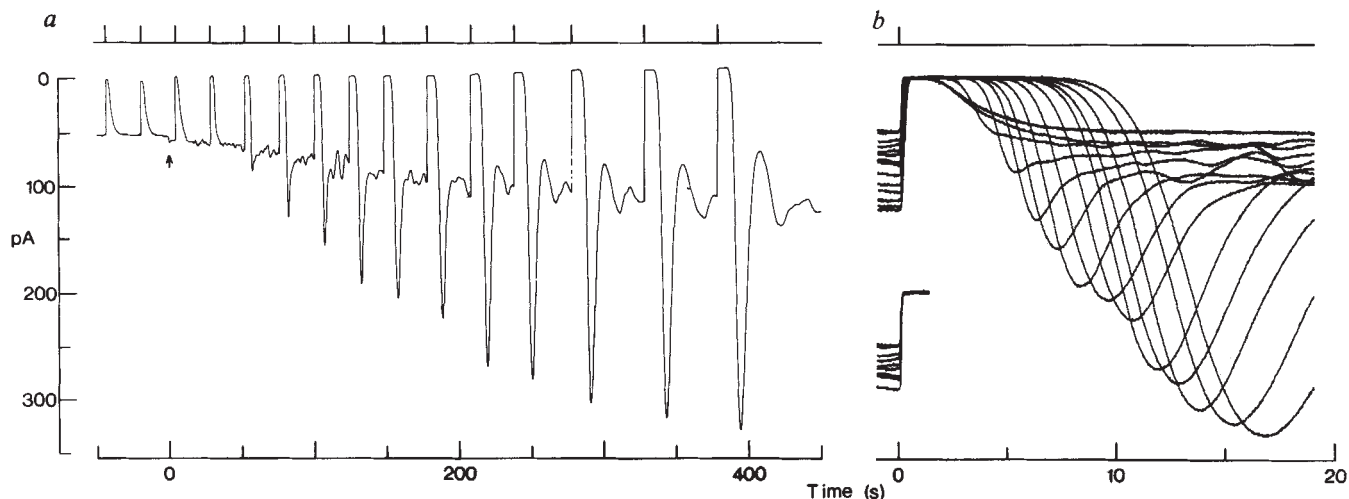


Fig. 1 Suction pipette current recorded from the outer segment of a rod during incorporation of calcium chelator into the cytoplasm by means of a patch pipette. *a*, Continuous record, with arrow at time zero indicating rupture of the patch and initiation of the whole-cell recording. *b*, Superimposed responses to bright flashes from *a*; inset, rising phases of first nine responses. The patch pipette solution contained (in mM): 92 K-aspartate, 7 NaCl, 1 NaHCO₃, 7.4 MgCl₂, 10 taurine, 1 Na₂-ATP, 1 Na₄-GTP and 10.4 BAPTA (BDH), buffered to pH 7.4 with KOH. The dissociation constants of BAPTA for Ca²⁺ and Mg²⁺ are 1.1×10^{-7} M and 17 mM, respectively; the dissociation time constant for Ca²⁺ is < 16 ms, and it is virtually unaffected by pH (see ref. 11); free concentrations of BAPTA, Mg²⁺ and Ca²⁺ were 8 mM, 5 mM and $\sim 10^{-9}$ M, respectively. The external bathing solution was normal Ringer's⁹. The patch-pipette voltage was clamped to -33 mV immediately after sealing against the cell, using a List EPC-7 patch-clamp amplifier. Patch pipette resistance was 3.2 M Ω ; seal resistance ≥ 1 G Ω ; whole-cell access resistance initially 15 M Ω , falling to 8 M Ω at 70 s; access time constant 200 μ s, giving cell capacitance 25 pF. The whole-cell current (not shown) was $\sim 60\%$ greater than the suction pipette current, as the suction pipette did not record the entire ROS current. Light stimuli (upper trace) delivered 207 photons μm^{-2} at 500 nm; 23.5°C. The prominent overshoot and damped oscillation may have resulted from a light-induced reduction in the free calcium level (see text).

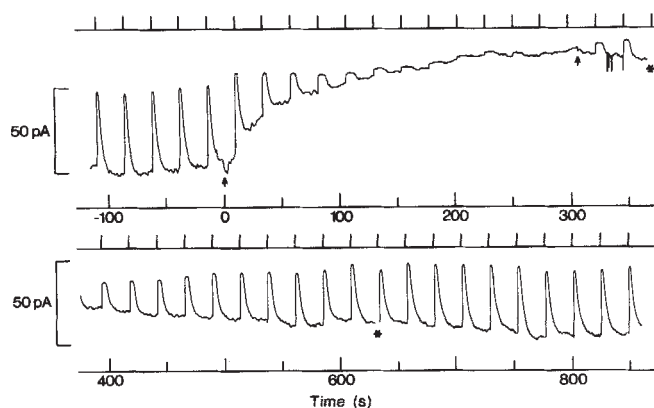


Fig. 2 Continuous record showing the decay of the 'dark current' on rupturing the membrane with a patch pipette containing 106 mM Na⁺ (at time 0), and recovery after successful withdrawal of the patch pipette (at 305 s). Na-aspartate (92 mM) replaced the usual 92 mM K-aspartate; 20 μ M BAPTA. Some drift occurred, particularly during the attempt to pull the pipette off (close to 200 s). The trace has been interrupted at the asterisks, which mark the times when the cell was repositioned in the suction pipette and the current trace was repositioned also. Clamp voltage -33 mV; whole-cell access resistance initially ~ 15 M Ω , later (after 50 s) ~ 40 M Ω . Flashes delivered 207 photons μm^{-2} at 500 nm; 23.8°C.

Figure 2 shows that the pipette contents exchanged with the cytoplasmic contents; here a patch pipette containing a high concentration (106 mM) of Na⁺ was used, and the first arrow indicates the time of patch rupture. The dark current then began declining with a time constant of ~ 50 s as might be expected for an increased internal sodium concentration, and the saturating responses increased in duration. After ~ 200 s an attempt was made to pull the patch pipette away from the cell, and this was successfully accomplished at 305 s (second arrow). At this point the cell immediately began recovering, presumably as it

pumped the Na⁺ out again. This recovery indicates that the decay of dark current was due not to damage but rather to diffusion of sodium into the cell. A similar time constant of decline of the dark current (45–65 s) was observed in four cells, and in two of these the pipette was pulled off without damaging the cell, which showed substantial recovery. In two other cells an intermediate concentration of Na⁺ (56 mM) was used, yielding a somewhat slower approach to a final dark current roughly 30% of its initial value; in both cases there was substantial recovery.

The following represent additional evidence that the substances introduced reach the outer segment: (1) pipettes containing BAPTA or cyclic GMP (see below) have a major effect on the ROS photocurrent; (2) using similar methods, A. Zimmerman and D. A. Baylor (unpublished) observed equilibration of the fluorescent dye carboxyfluorescein throughout the entire rod in about 2 min; (3) a few experiments in which the inner segment was drawn into the suction pipette and the patch pipette was attached to the outer segment gave similar results, although it seemed much more difficult to rupture the patch successfully.

Hagins and Yoshikami¹² reported that incorporation of small quantities of calcium buffers into rods by a vesicle fusion technique led to 'buffering' of the response to light, that is, a brighter flash was required to elicit a given fractional response amplitude. We tested the flash sensitivity in the presence of BAPTA and observed the opposite result—the buffer increased the sensitivity to flashes. The lowest trace in Fig. 3a was obtained before rupturing the membrane patch, whereas the upper three traces were obtained with the same dim flashes at the indicated times after obtaining the whole-cell recording. Incorporation of 10 mM buffer increased greatly the amplitude of the responses to dim flashes, slowed their time course, and produced an overshoot in the recovery phase similar to that observed with bright flashes. Similar increases in flash sensitivity were obtained for seven whole-cell recordings with 2 or 10 mM BAPTA in the pipette. Figure 3b confirms that with only a trace amount (20 μ M) of BAPTA in the whole-cell pipette, very little change in sensitivity occurred, over a period of at least 8 min.

Our interpretation of these results is that the calcium buffer reaches the ROS, and that despite this the saturating response to light remains, with an unchanged rate of rise, and the flash sensitivity increases dramatically. These observations seem to be contrary to the calcium hypothesis of transduction in rods, but one major qualification is that we have no evidence that the free calcium concentration in the outer segment is even approximately equal to that in the patch pipette. We expect that the cell probably possesses powerful mechanisms for adjusting the

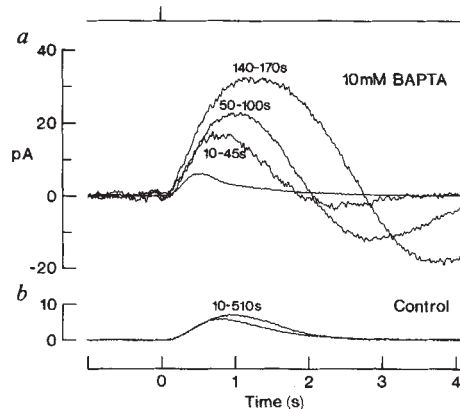


Fig. 3 Flash sensitivity in the presence of BAPTA and in control conditions. *a*, Patch pipette solution and experimental method as in Fig. 1 (10.4 mM BAPTA, no added Ca^{2+}). The lowest trace is the average of 39 responses obtained before rupturing the patch (dark current ~ 40 pA), and the upper traces represent 3–5 responses obtained at the indicated times after rupture (mean dark current ~ 50 pA; 23.6°C). The response amplitude was roughly linear with flash intensity, and the superlinearity found with reduced external calcium^{16,24,25} was not observed. The increased flash sensitivity in the presence of BAPTA probably accounts for the observation in Fig. 1 of an increased duration of bright flash responses. *b*, Control with only $20\ \mu\text{M}$ BAPTA (to bind stray Ca^{2+}) and $5\ \text{mM}$ MgCl_2 , but otherwise as in Fig. 1. Lower trace, 57 flashes before rupture; upper trace, 42 flashes after rupture of the patch; mean dark current 46 and 38 pA respectively; 23.7°C . In *a* and *b* dim flashes delivered 2.3 and 1.2 photons μm^{-2} , respectively, at 500 nm; clamp voltage $-33\ \text{mV}$.

mean level of free calcium in the long term, so that in our experiments the resting dark concentration of free calcium may have been near normal. The fact that the presence of exogenous calcium buffer had a pronounced effect on the ROS photocurrent (for example, the overshoot) indicates that the buffer was operating within its buffering range, that is, it was neither fully bound nor fully unbound. Thus, the continued suppression of the dark current by bright flashes, together with the increased flash sensitivity, still seem to be inconsistent with the calcium hypothesis. We are not suggesting that elevated internal calcium levels in the intact cell are unable to lead to closure of sodium channels, but rather that the normal light response does not appear to operate in this way. Indeed, if calcium leaks into the cytoplasm through the light-sensitive channels in the dark^{13–15} and is removed by $\text{Na}^+/\text{Ca}^{2+}$ exchange^{13–16,26}, cytoplasmic calcium levels might actually drop in the light; this could explain the pronounced overshoot observed with responses to bright flashes (Fig. 1). Evidence in support of such a reduction of calcium in the light has been reported recently¹⁷.

Figure 4 shows a result obtained with $2\ \text{mM}$ cyclic GMP in the patch pipette. Shortly after rupturing the patch (second arrow) the dark current began increasing dramatically, but bright flashes continued to suppress the elevated current. Over a period of a few minutes the rising phases of the responses (Fig. 4*b*) slowed somewhat, and the responses remained saturated for a very long time. Similar results have been obtained recently using the same method¹⁸, and in previous experiments on injection of cyclic GMP^{19–21} or perfusion with phosphodiesterase inhibitors²². Although cyclic GMP incorporated in the dark elicited very large currents, we found that in the presence of a moderate steady background light, cyclic GMP had virtually no effect (not shown). While the rapid increase in dark current is qualitatively what would be expected if cyclic GMP were the internal messenger, the substantial increase in the time for which the light-sensitive channels remain closed following a bright flash is not what would be expected based on the simplest model of light-activated hydrolysis of cyclic GMP.

A preliminary report of some of this work has appeared elsewhere²³. The work was supported by grants from the MRC, by a Royal Society Locke Research Fellowship, and by an EMBO Fellowship. We thank Dr T. J. Rink for helpful comments on the manuscript and acknowledge the 1983 Cold Spring Harbor course, where the patch technique was learnt.

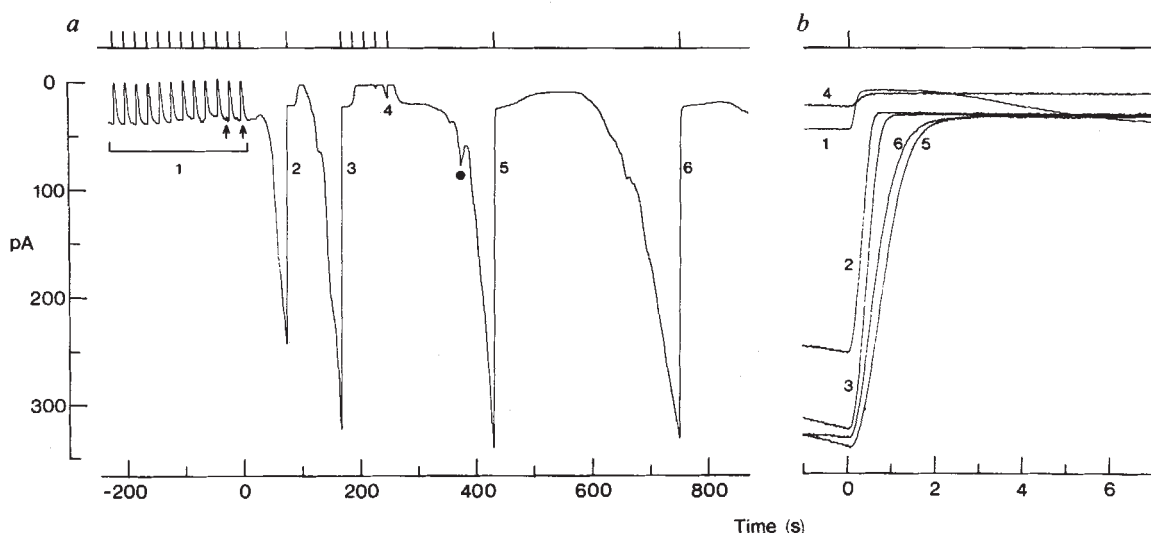


Fig. 4 Rod outer segment suction pipette current recorded with $2\ \text{mM}$ cyclic GMP in the patch pipette. *a*, Pipette sealed at first arrow, and patch ruptured at second arrow. Bright flashes (shown at the top) were stopped at this point. After a delay the dark current increased dramatically, and occasional light flashes were delivered. For the flashes labelled 2, 3, 5 and 6 the current did not fall immediately to zero, possibly because of a light-insensitive $\text{Na}^+/\text{Ca}^{2+}$ exchange current¹⁵ induced by the presumed influx of Ca^{2+} . ●, Rod was repositioned slightly in the suction pipette. *b*, Superimposed bright flash responses from *a*. The patch pipette solution contained $2\ \text{mM}$ cyclic GMP, but was otherwise as in Fig. 2*b*, except for the absence of ATP and GTP. Light flashes delivered 108 photons μm^{-2} at 500 nm.

Received 17 October; accepted 14 December 1984.

1. Miller, W. H. *Curr. Topics Membranes Transport* 15 (1981).
2. Korenbrot, J. I. *Progress in Retinal Research* Vol. 4 (Pergamon, Oxford, in the press).
3. Hagins, W. A. *Rev. Biophys. Bioengng* 1, 131-158 (1972).
4. Yoshikami, S. & Hagins, W. A. in *Biochemistry and Physiology of Visual Pigments* (ed. Langer, H.) 245-255 (Springer, Berlin, 1973).
5. Hubbel, W. L. & Bownds, M. D. *A. Rev. Neurosci.* 2, 17-34 (1979).
6. Woodruff, M. L. & Bownds, M. D. *J. gen. Physiol.* 73, 629-653 (1979).
7. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Eur. J. Physiol.* 391, 85-100 (1981).
8. Yau, K.-W., Lamb, T. D. & Baylor, D. A. *Nature* 269, 78-80 (1977).
9. Baylor, D. A., Lamb, T. D. & Yau, K.-W. *J. Physiol., Lond.* 288, 589-611 (1979).
10. Robinson, P. R., MacLeish, P. R. & Lisman, J. Soc. Neurosci. Abstr. 9, 165 (1983).
11. Tsien, R. Y. *Biochemistry* 19, 2396-2404 (1980).
12. Hagins, W. A. & Yoshikami, S. in *Vertebrate Photoreception* (eds Barlow, H. B. & Fatt, P.) 97-139 (Academic, London, 1977).
13. Yau, K.-W. & Nakatani, K. *Nature* 309, 352-354 (1984).
14. Hodgkin, A. L., McNaughton, P. A., Nunn, B. J. & Yau, K.-W. *J. Physiol., Lond.* 350, 649-680 (1984).
15. Yau, K.-W. & Nakatani, K. *Nature* 311, 661-663 (1984).
16. Bastian, B. & Fain, G. L. *J. Physiol., Lond.* 330, 307-329 (1982).
17. Yau, K.-W. & Nakatani, K. *Nature* 313, 579-582 (1985).
18. Cobbs, W. H. & Pugh, E. N. Jr *Nature* 313, 585-587 (1985).
19. Miller, W. H. *J. gen. Physiol.* 80, 103-123 (1982).
20. Kawamura, S. & Murakami, M. *Jap. J. Physiol.* 33, 789-800 (1983).
21. MacLeish, P. R., Schwartz, E. A. & Tachibana, M. *J. Physiol., Lond.* 348, 645-664 (1984).
22. Capovilla, M., Cervetto, L. & Torre, V. J. *J. Physiol., Lond.* 343, 277-294 (1983).
23. Lamb, T. D., Matthews, H. R. & Torre, V. J. *J. Physiol., Lond.* 360 (in the press).
24. Yau, K. W., McNaughton, P. A. & Hodgkin, A. L. *Nature* 292, 502-505 (1981).
25. Owen, W. G. & Torre, V. *Biophys. J.* 41, 305-324 (1983).
26. Gold, G. H. & Korenbrot, J. I. *Curr. Topics Membranes Transport* 15, 307-330 (1981).

Cyclic GMP can increase rod outer-segment light-sensitive current 10-fold without delay of excitation

W. H. Cobbs & E. N. Pugh Jr

Departments of Neurology and Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

To test the hypothesis that cyclic GMP is the internal messenger coupling rhodopsin activation to membrane excitation in vertebrate rod photoreceptors, we used a novel technique combining measurement of membrane currents of isolated salamander rods with a suction-electrode and the introduction of cyclic GMP through a whole-cell recording patch pipette. Rupture of an attached patch was followed by a rapid (~ 10 s), approximately 10-fold increase in outer-segment membrane current, all of which was light-sensitive. There was little change in the rising phase of the response to a saturating flash, but the duration of the saturated phase of the response increased ~ 10 -fold. The effects reversed completely within 3-4 min after withdrawal of the cyclic GMP-containing patch pipette. A formal kinetic analysis shows that the first two observations are inconsistent with the postulate that cyclic GMP opens the light-sensitive conductance by simple binding to channels, unless free cyclic GMP in the outer segment is assumed to be much lower than published estimates, and most of the outer-segment cyclic GMP is bound and inexchangeable on the timescale of 200 ms. Furthermore, our results suggest that rod cyclic GMP is not involved solely in keeping the light-sensitive conductance open, but may also affect the activity of the phosphodiesterase that mediates cyclic GMP hydrolysis.

Isolated rods from larval tiger salamander retina were secured in a recording suction electrode¹, and a gigaseal patch micropipette containing cyclic GMP was positioned against the exposed inner-segment membrane (see Fig. 1B). Recordings were made of both the suction- and patch-electrode currents (whole-cell voltage-clamp mode). Figure 1A shows the outer-segment membrane current recorded with a suction electrode in these conditions. Initially, at time point labelled (1) in Fig. 1A, a saturating flash of 1,000 rhodopsin isomerizations produced a brief 52-pA photocurrent, which decayed to the resting value in less than 10 s. When the patch was ruptured (at WC in Fig. 1A), the outer-segment membrane current began increasing

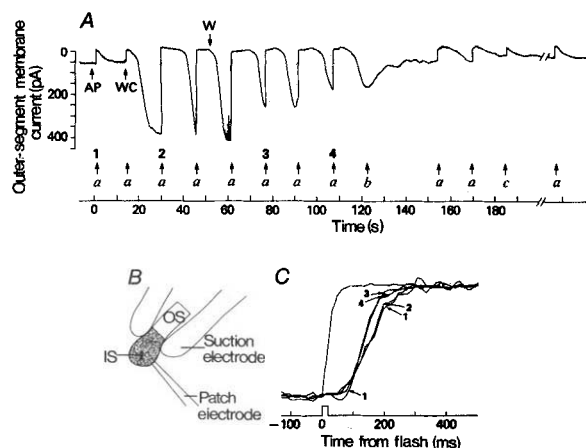


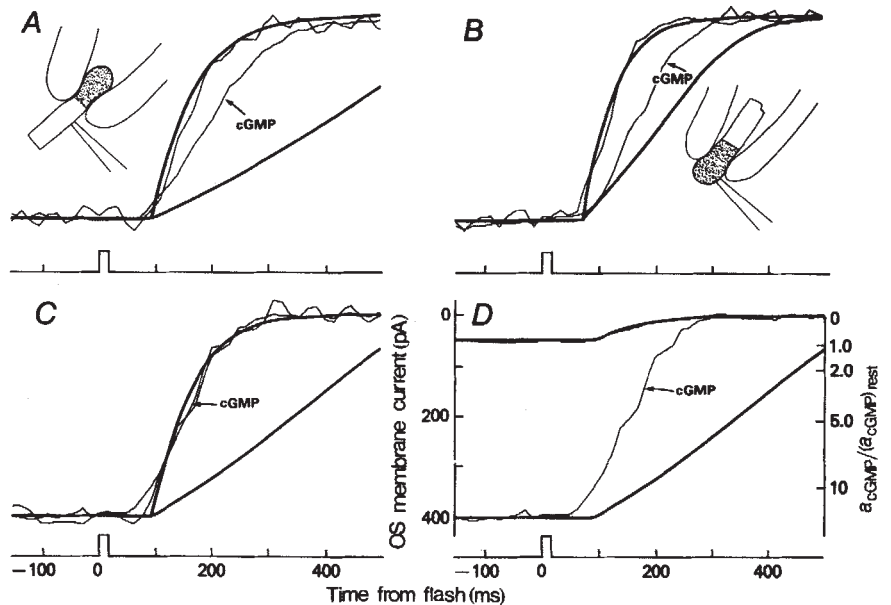
Fig. 1 Effect of exogenous cyclic GMP on salamander rod dark- and photocurrents. **A**, Membrane current of salamander rod outer segment recorded at 30 Hz bandwidth with a suction electrode in the configuration shown in **B**. At 'AP' a ~ 2 G Ω attached 'patch' was formed on the inner segment with a 20-M Ω Kovar sealing glass electrode. At 'WC' the patch was ruptured with negative pressure and the dark current underwent a rapid, large increase. At 'W' the patching electrode was successfully detached. Arrows below the trace show the times of 20-ms flashes: *a* indicates a flash large enough to cause 1,000 isomerizations, *b* 100 and *c* 10; 1-4 indicate responses shown enlarged in **C**. **B**, Line drawing traced from a videotape record of this experiment, to show the position of the patch electrode and of the cell in the suction electrode. Rod outer segment (OS) diameter, 12 μ m. IS, inner segment. **C**, The rising portions of the photocurrents labelled 1-4 in **A**, normalized to maximum amplitude. The leftmost curve shows the step-response of the complete recording system (amplifier, multipole filter, FM tape recorder, computer A/D) to a current step initiated at the same point as the stimulus flash.

Methods. The patch electrode contained 100 mM K-acetate, 2 mM MgCl₂, 10 mM HEPES, 5 mM NaCl, 50 μ M EGTA and 5 mM cyclic GMP, adjusted to pH 7.5 with KOH. The cell was bathed in Ringer's solution containing 110 mM NaCl, 2.5 mM KCl, 1 mM CaCl₂, 1 mM NaH₂PO₄, saturated with 100% oxygen at 23 °C. All experiments were monitored and videotaped with an infrared ($\lambda > 820$ nm) television system.

rapidly, becoming eight-fold greater within 10 s. A light flash of the same intensity (at (2) in Fig. 1A) then caused a 400-pA photocurrent. This saturated response reached the same level as the response before rupture of the patch, indicating that the entire increase in outer-segment membrane (dark) current was light sensitive. The duration of the saturated phase of the photocurrent responses to flashes also increased ~ 10 -fold. After withdrawal of the cyclic GMP-containing pipette (at W in Fig. 1A) the membrane current gradually diminished to baseline level and photocurrent responses returned to normal over a period of 3-4 min, demonstrating that the effects of exposure to exogenous cyclic GMP are reversible. Normalizing the photocurrents by maximum amplitude emphasizes that the rising phase 10 s after rupture hardly differs from that observed before rupture, despite an eight-fold increase in current magnitude (Fig. 1C). Control experiments using 10- μ s flashes of up to 10^7 isomerizations showed that rods can generate photocurrents that match the recording system step response (Fig. 1C, leftmost curve), so the lack of change in photocurrent rising phase is due neither to instrumental limitation nor to cell cable properties.

We have observed similar effects of cyclic GMP in 15 other cells. The most striking effect of cyclic GMP in all cases was a large, rapid increase in dark- and photocurrents, as expected from previous microelectrode studies²⁻⁶. The current often increased beyond the range of the recording system (~ 600 pA from baseline). When the patch pipette contained no cyclic GMP, cells maintained normal dark- and photocurrents for up to 30 min.

Fig. 2 Comparison of photocurrent responses of four different rods before and after rupture of the membrane patch, with theoretical predictions for the delay in rising phase in the presence of exogenous cyclic GMP. **A–C**, Thin curves show photocurrents recorded before and after patch rupture and the thick lines the theoretical predictions for the response in the presence of exogenous cyclic GMP in normalized amplitude coordinates, to facilitate the evaluation of the fit to the pre-cyclic GMP photocurrent. **A**, One of two rods in which cyclic GMP was introduced directly into the outer segment. The baseline photocurrent recorded by the suction electrode was 20 pA; 16 s after patch rupture the photocurrent ('cGMP') recorded with the voltage-clamp electrode was 400 pA. **B**, The baseline photocurrent was 60 pA; at $t = 6$ s after patch rupture, the photocurrent ('cGMP') measured 320 pA. **C**, The baseline response was 52 pA; at 10 s after patch rupture, the photocurrent was 400 pA (data from Fig. 1C). **D**, The responses (thin curves) to flashes 1 and 2 of the cell shown in Fig. 1 plotted using an absolute scale (left-hand ordinate) for outer-segment (OS) membrane current. The thick line drawn through the upper (pre-cyclic GMP) response was calculated by integrating the Michaelis rate expression, $dS/dt = -V_{\max}S/(S + K_m)$, in which $S = a_{\text{cGMP}}$, $S_0 = 40 \mu\text{M}^{10}$, $k_m = 80 \mu\text{M}^9$, and $V_{\max}$ was adjusted for optimal visual fit; $S(t)$ was then converted to normalized photocurrent Y , with the inverse of equation (2): $Y(t) = S(t)/(S(t) + K_D)$, with $K_D = 730 \mu\text{M}$. In this fitting procedure we approximated the multi-order delay of the very early rising phase with a pure delay that appears in the theoretical (thick) curves as absolute dead time after the flash before the curve begins to ascend. The value of $K_D = 730 \mu\text{M}$ is required by the assumptions that $(a_{\text{cGMP}})_{\text{rest}}$, the baseline free cyclic GMP in a cell with a normal 50-pA response, is $40 \mu\text{M}^{10}$, and that maximal opening of the outer-segment membrane conductance results in a dark current of 1 nA. The lower thick curve is the response predicted in the presence of exogenous cyclic GMP, computed with the same values of k_m and $V_{\max}$, but with the pre-flash $S_0 = 487 \mu\text{M}$, the value required by equation (2) to account for the 400-pA dark current after patch rupture, immediately before the flash. The values of $V_{\max}$ estimated by fitting the theoretical curves to the pre-cyclic GMP baseline responses in **A–C** are 1,100, 1,700 and 1,600 $\mu\text{M s}^{-1}$, respectively. For a salamander rod outer segment with a 1.5-pl water space, a PDE velocity of 1,000 $\mu\text{M s}^{-1}$ in response to a 1,000-isomerization flash corresponds to an intrinsic enzyme activity of 10^6 cyclic GMP molecules hydrolysed per s per isomerization. All cells showed increased duration of the saturated phase of the photocurrent after exposure to exogenous cyclic GMP. We took the interval from the time at which the photocurrent reached saturation until the time it recovered 10% of its pre-flash dark current as a measure of the saturated phase duration. For the pre-cyclic GMP and post-cyclic GMP responses in **A–C**, the durations (s) of the saturated phases were: **A**, pre, 1.0/post, 11.0; **B**, pre, 1.1/post, 5.0; **C**, pre, 1.0/post, 8.8.

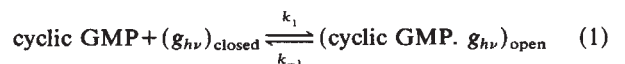


Two of the 16 cells examined here were penetrated successfully through the outer segment, the inner segment being held in the suction electrode. This provided an opportunity to assess the performance of our voltage- and space-clamp for large outer-segment membrane currents, because the inner segment contains only voltage-dependent, not light-dependent, conductances. In both cases, the inner-segment photocurrent measured with the suction electrode was less than 5% of that measured with the voltage-clamp electrode. On symmetrical grounds, we assume that cells penetrated through the inner segment were reasonably well voltage- and space-clamped despite large cyclic GMP-induced membrane currents.

Figure 2A, B (thin curves) show the rising phases of the photocurrents before and after patch rupture for two more cells, highlighting the fact that cyclic GMP can increase greatly the outer-segment membrane current without causing large changes in photocurrent rising phase.

We were surprised that cyclic GMP can cause large increases in outer-segment light-sensitive membrane current with little or no retardation of the photocurrent rising phase. Previous investigations using different techniques found significant slowing of the rising phase of photovoltages after cyclic GMP injections^{2–4}. This slowing was interpreted as the extra time required for cyclic GMP phosphodiesterase (PDE) to hydrolyse the injected cyclic GMP⁵. Therefore, we have examined the quantitative predictions of a model in which a decrease in the cytoplasmic activity of cyclic GMP (a_{cGMP} is the excitatory message to the light-sensitive conductance. This model postulates that (1) normally in the dark a balance of guanylate cyclase and low PDE activity maintains a_{cGMP} at a constant level; (2) light causes a_{cGMP} to fall by increasing PDE activity, with a time course independent

of a_{cGMP} ; (3) a_{cGMP} regulates the outer-segment light-sensitive conductance ($g_{\text{h}\nu}$), through a simple binding reaction:



Under reasonable assumptions⁷, equation (1) predicts a hyperbolic relationship between a_{cGMP} and the fraction of the conductance open at steady state. Moreover, the speed of the response to intense flashes⁸ requires the fraction of channels open to equilibrate with any change in a_{cGMP} in a few milliseconds. Thus, on a timescale of tens of milliseconds, the following relation always holds:

$$a_{\text{cGMP}} = [Y/(1 - Y)]K_D \quad (2)$$

where Y is the fraction of outer-segment light-sensitive conductance open, and $K_D = k_{-1}/k_1$, the dissociation constant of reaction (1). We used this relation to predict photocurrent rising phases expected in the presence of exogenous cyclic GMP.

Figure 2 compares predictions of this model with our observations. The thin curves are the photocurrents before and after patch rupture. The upper thick curve is a fit of the model to the control (pre-cyclic GMP) response and the lower thick curve is the prediction for the post-patch rupture photocurrent. In each case, theory predicts that exogenous cyclic GMP should slow the photocurrent rising phase far more than observed.

The simplified model neglects diffusion of cyclic GMP into the outer segment from the patch pipette. Several arguments support the assumption that the rate of cyclic GMP diffusion into the outer segment from the inner segment after patch rupture is at least two orders of magnitude slower than its apparent rate

of destruction during the ~200-ms photocurrent rising phase. First, it takes at least 10 s for the dark current to reach its large steady-state value after patch rupture. Second, return to the initial baseline value after removal of the patch electrode containing cyclic GMP takes several minutes, even with repeated saturating flashes that should reduce the outer-segment a_{cGMP} to a very low value. Third, A. Zimmerman and D. A. Baylor (personal communication) have found, using the preparation and geometry shown in Fig. 1B, that it takes ~2 min for the outer segment to fill with the hydrophilic dye carboxyfluorescein, when it is introduced from the patch pipette.

Why does the model fail to account for the rising phase data? Adjusting the numerical values of the parameters within limits set by published results^{9,10} does not produce satisfactory improvement. In our experiments, however, it may be that $a_{\text{cGMP}} \ll k_m$ of PDE (P. A. Liebman, personal communication), which would yield approximate rising phase invariance. To obtain $a_{\text{cGMP}} \ll k_m$, either $(a_{\text{cGMP}})_{\text{rest}}$ must be lower or k_m higher than the published estimates [$k_m = 80 \mu\text{M}$ (ref. 9); $a_{\text{cGMP}} = 40 \mu\text{M}$ (ref. 10)]. Setting k_m of PDE to 1 mM, without concomitant alteration of $(a_{\text{cGMP}})_{\text{rest}}$, requires $V_{\text{max}} = 15,000 \mu\text{M s}^{-1}$, a specific PDE activity more than two orders of magnitudes higher than any published value for light-activated cyclic GMP hydrolysis by disk membranes in the presence of intracellular amounts of ATP and GTP. Alternatively, one can adjust $(a_{\text{cGMP}})_{\text{rest}}$ to 1 μM , without changing k_m , but this requires that over 90% of the total outer-segment cyclic GMP be bound and not exchangeable during the response rising phase. It also requires that in our experiments the free cyclic GMP was elevated from 1 μM to ~10–15 μM 10 s after patch rupture, which is not implausible, given the approximate equilibration time of 2 min found by Zimmerman and Baylor (personal communication). The notion that most of the outer-segment cyclic GMP is bound, and not exchangeable on the timescale of the rising phase of the response, can explain the puzzling finding that even intense flashes do not cause a rapid decrease in total outer-segment cyclic GMP, measured in preparations quickly quenched after flash exposure^{11,12}.

A more drastic modification of the model assumes cooperative binding of cyclic GMP to the channels, rather than the non-cooperative binding of equation (1). This gives a steeper relation between a_{cGMP} and $g_{\text{h}\nu}$, so that a small increment of a_{cGMP} above its resting value would move $g_{\text{h}\nu}$ to a point nearly half-way up the binding curve. The PDE activity produced by the light flash could then hydrolyse both native and exogenous cyclic GMP in approximately the same time. Calculations incorporating cooperative binding but no change in $(a_{\text{cGMP}})_{\text{rest}}$ show improvement in fitting our rising phase data.

It has been proposed that cyclic GMP controls the light-sensitive conductance through a cyclic GMP-dependent protein kinase^{13,14}. Reconciling the large increase in dark current with a small change in photocurrent rising phase requires that protein kinase activity depends steeply on a_{cGMP} . Inclusion of protein kinase and phosphoprotein phosphatase as part of the transduction mechanism does not, however, provide an explanation of our rising phase data and moreover would impose a significant energy cost on the cell¹⁵. Dephosphorylation of each open conductance must occur every millisecond to enable the rising phase of the photocurrent in response to an intense flash to be as short as 700 μs ⁸. If the unitary current¹⁶ is 2 fA and the normal dark current 50 pA, then the kinase/phosphatase reactions would require 2.5×10^7 ATP molecules s^{-1} to maintain the normal dark current, close to the 10^8 ATP molecules s^{-1} required by the Na/K exchanger to maintain the resting Na gradient for a 50-pA dark current.

To summarize, the relative invariance of the rising phase of photocurrents before and soon after patch rupture provides a strong constraint for the mechanism by which cyclic controls the outer-segment light-sensitive conductance. The simplest explanation of this observation is that the free cyclic GMP in the outer segment is only a small fraction of the total, and that bound cyclic GMP does not exchange within 200 ms. The strik-

ing prolongation of the photocurrent following exposure to exogenous cyclic GMP (Fig. 1A), however, cannot be explained within the framework of the model stated above. Rather it requires that the time course of PDE activity be affected by previous exposure to excess cyclic GMP, or by some agent such as calcium that enters through the light-sensitive conductance. This conclusion follows because prolonged dark current suppression occurs when a_{cGMP} can no longer be elevated.

Other transduction theories postulate that change in a_{Ca} , not a_{cGMP} , is the excitatory message for phototransduction^{17,18}. The large increase in dark current caused by exogenous cyclic GMP could be explained by an increase in cyclic GMP-dependent calcium pumping¹⁹, which would lower a_{Ca} in the rod cytoplasm. Little change in photocurrent rising phase should occur after exposure to exogenous cyclic GMP, if the same amount of calcium were released per isomerization and the number of calcium ions released greatly exceeded the number of conductance sites.

The recent results of Fesenko *et al.*²⁰, brought to our attention during review of the present paper, provide important constraints on the form of the binding relation. In view of their data, equation (1) would have to be modified to

$$n \text{ cyclic GMP} + (g_{\text{h}\nu})_{\text{closed}} \xrightleftharpoons[k_{-1}]{k_1} [\text{in (cyclic GMP).} g_{\text{h}\nu}]_{\text{open}} \quad (1')$$

and equation (2) would then become

$$a_{\text{cGMP}} = [Y/(1-Y)]^{1/n} \cdot K'_D \quad (2')$$

where $K'_D = (k_{-1}/k_1)^{1/n}$ is the value of a_{cGMP} at which $g_{\text{h}\nu}$ is half-open, and $n=2$. Refitting our data on the photocurrent rising phase with these modified equations, we found a notable improvement over the non-cooperative binding model when all other parameters were unchanged. However, to keep $(a_{\text{cGMP}})_{\text{rest}}$ equal to 40 μM , K'_D in equation (2') had to be 171 μM , at variance with the direct estimate of 30 μM made by Fesenko *et al.*²⁰. If cyclic GMP is the only agent that regulates the outer-segment light-sensitive conductance, then equation (2'), and the assumption that 95% of the outer-segment conductance is shut off in the resting (dark) state, require $(a_{\text{cGMP}})_{\text{rest}}$ to be 7 μM . Refitting our rising phase data by forcing $(a_{\text{cGMP}})_{\text{rest}}$ to be 7 μM produces good agreement for k_m values as low as 80 μM . We emphasize, however, that the results of Fesenko *et al.* provide no insight into the lengthening of the saturated phase of the photocurrent after exposure to exogenous cyclic GMP, which has to be explained in terms of cyclic GMP-dependent biochemical events prior to the binding relation, equation (2').

This work was supported by USPHS grants EY-02660 and EY-01583. We acknowledge consultations with Drs A. Zimmerman, T. D. Lamb and D. A. Baylor in the development of this preparation; they have also observed large increases of outer-segment membrane current in the presence of exogenous cyclic GMP. We thank Drs B. Rosner, P. Mueller and P. A. Liebman for helpful comments.

Received 19 November 1984; accepted 4 January 1985.

1. Baylor, D. A., Lamb, T. D. & Yau, K.-W. *J. Physiol., Lond.* **288**, 589–611 (1979).
2. Nicol, G. D. & Miller, W. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5217–5220 (1978).
3. Miller, W. H. & Nicol, G. D. *Nature* **280**, 64–66 (1979).
4. Miller, W. H. *J. gen. Physiol.* **80**, 103–123 (1983).
5. Waloga, G. *J. Physiol., Lond.* **341**, 341–357 (1983).
6. MacLeish, P. R., Schwartz, E. A. & Tachibana, M. *J. Physiol., Lond.* **348**, 645–664 (1984).
7. Baylor, D. A., Hodgkin, A. L. & Lamb, T. D. *J. Physiol., Lond.* **242**, 685–727 (1973).
8. Penn, R. D. & Hagins, W. A. *Biophys. J.* **12**, 1073–1094 (1972).
9. Yee, R. & Liebman, P. A. *J. biol. Chem.* **253**, 8902–8909 (1978).
10. Woodruff, M. L. & Bownds, M. D. *J. gen. Physiol.* **73**, 629–653 (1979).
11. Kilbride, P. & Ebrey, T. G. *J. gen. Physiol.* **74**, 415–426 (1979).
12. Govardovskii, V. I. & Berman, A. L. *Biophys. Struct. Mechanism* **7**, 125–130 (1981).
13. Farber, D. B., Brown, B. M. & Lolley, R. N. *Vision Res.* **18**, 479–492 (1978).
14. Liebman, P. A. & Pugh, E. N. Jr. *Curr. Topics Membranes Transp.* **15**, 157–170 (1981).
15. Lamb, T. D. in *Recent Advances in Physiology* (ed. Baker, P. F.) 29–65 (Churchill, London, 1984).
16. Detwiler, P. B., Conner, J. D. & Bodoia, R. D. *Nature* **300**, 59–61 (1982).
17. Yoshikami, S. & Hagins, W. A. *Biophys. J.* **10**, 60a (1970).
18. Liebman, P. A., Mueller, P. & Pugh, E. N. Jr. *J. Physiol., Lond.* **347**, 85–110 (1984).
19. George, J. S. & Hagins, W. A. *Nature* **303**, 344–348 (1983).
20. Fesenko, E. E., Kolesnikov, S. S. & Arkady, A. L. *Nature* **313**, 310–313 (1985).

Na channels in skeletal muscle concentrated near the neuromuscular junction

Kurt G. Beam*, John H. Caldwell†
& Donald T. Campbell*‡

* Department of Physiology and Biophysics, University of Iowa, School of Medicine, Iowa City, Iowa 52242, USA

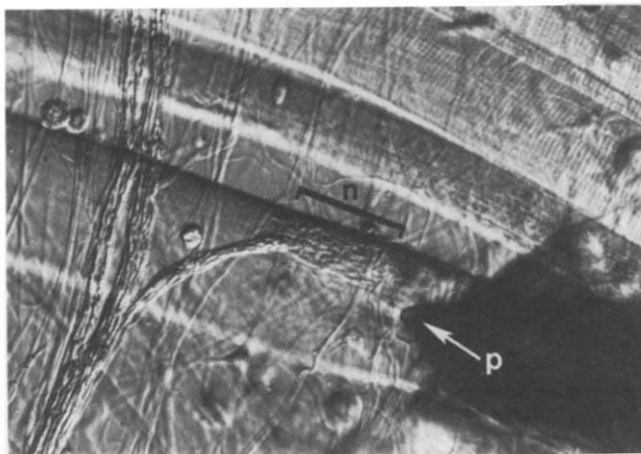
† Department of Molecular and Cellular Biology, National Jewish Hospital and Research Center, Denver, Colorado 80206 and Department of Physiology, University of Colorado Medical School, Denver, Colorado 80262, USA

Neuronal function depends crucially on the spatial segregation of specific membrane proteins, particularly the segregation associated with sites of synaptic contact. Understanding the factors governing this localization of proteins is a major goal of cellular neurobiology. A conspicuous example of synaptic specialization is the almost exclusive localization of vertebrate skeletal muscle acetylcholine (ACh) receptors to the subsynaptic membrane of the neuromuscular junction (for example, refs 1, 2). The localization of other membrane proteins in skeletal muscle has been much less studied, but a knowledge of their distribution is crucial for understanding the factors governing regional specialization. We have explored the distribution in muscle of the voltage-gated Na channel responsible for the action potential using the loose patch-clamp technique³⁻⁵, and have measured Na currents in 5–10 μm -diameter membrane patches as a function of distance from the end plate region of snake and rat muscle fibres. Here we report that the Na current density immediately adjacent to the endplate is 5–10-fold higher than at regions away from the endplate. The increased Na current density falls off rapidly with distance, reaching the background level 100–200 μm from the endplate. Although one might expect ACh receptors to be concentrated near the region of ACh release, such a concentration for Na channels, which propagate the impulse throughout the length of the cell, is surprising and suggests that factors similar to those responsible for concentrating ACh receptors at the endplate also operate to concentrate Na channels.

The loose patch voltage-clamp technique allows detailed mapping of Na current densities. In this method a fire-polished pipette is pressed against the surface of the muscle fibre, which partially electrically isolates a 5–10- μm -diameter patch of membrane and current flowing through the patch is measured⁴ in response to steps of membrane potential. To map Na current densities it is necessary to visualize the endplate clearly and to follow a single fibre over a considerable length. Therefore, we have used twitch fibres from the single-fibre-thick obliquus externus abdominis muscle of the garter snake (*Thamnophis s. sirtalis*) and enzymatically dissociated fibres¹ from the flexor digitorum brevis (FDB) muscle of the rat (Fig. 1). The position of recording sites relative to the endplate was measured typically from photomicrographs similar to those in Fig. 1.

Currents were recorded from each patch for a sequence of test pulses of increasing amplitude (see Fig. 2 inset). Figure 2 also shows the peak Na current obtained from such sequences recorded at varying distances from the endplate regions of snake and rat muscle fibres. In both snake and rat, the currents recorded near the endplate are much larger than those recorded far away. Similar results were obtained in all fibres studied in detail (five snake and four rat fibres). In the snake fibres, the ratio of peak Na current density at sites $\leq 40 \mu\text{m}$ from the endplate to Na current density at sites $\geq 245 \mu\text{m}$ away was 5.5 ± 2.4 (mean $\pm$ s.d., range 1.8–7.7, total of 14 'near' sites and 16 'far' sites); in four of these fibres the largest near-endplate density was 5.0–6.4 times larger than the largest distal density,

a



b

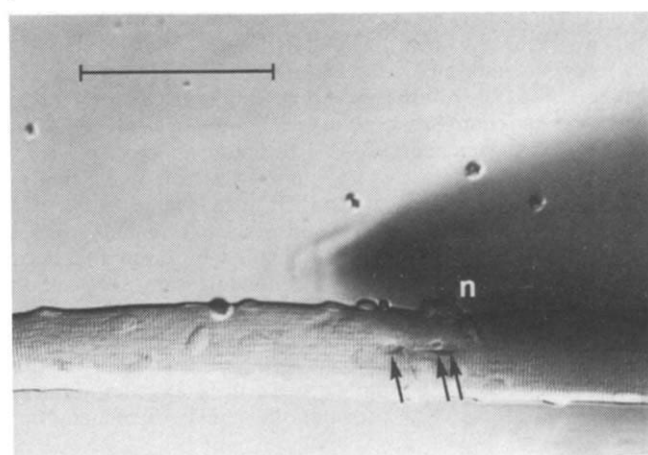


Fig. 1 Photomicrographs of the snake and rat muscle preparations viewed with a Nikon Diaphot equipped with Hoffman modulation contrast optics to facilitate visualization of endplates. *a*, Obliquus externus abdominis muscle of the garter snake (*Thamnophis s. sirtalis*). Clearly visible are the orifice (p) of the loose patch pipette pressed against the surface of a muscle fibre and the terminal (n) of the nerve innervating that fibre. *b*, Single fibre dissociated¹ enzymatically from rat FDB muscle. The remnants of the nerve terminal (n) remained attached to this fibre. The shadow above the fibre is the pipette tip, moved out of the plane of focus. Typically, a visible 'dimple' remained in dissociated fibres after the tip of the pipette had been removed from the surface of the fibre. Arrows indicate the dimples remaining at three separate recording positions of the pipette. Currents were recorded from the spot associated with the rightmost dimple, which has begun to fade, about 70 min before the photograph was taken. Currents were recorded from the spots associated with the other two dimples a few minutes before the photograph was taken. Scale bar, 100 μm (*a*, *b*).

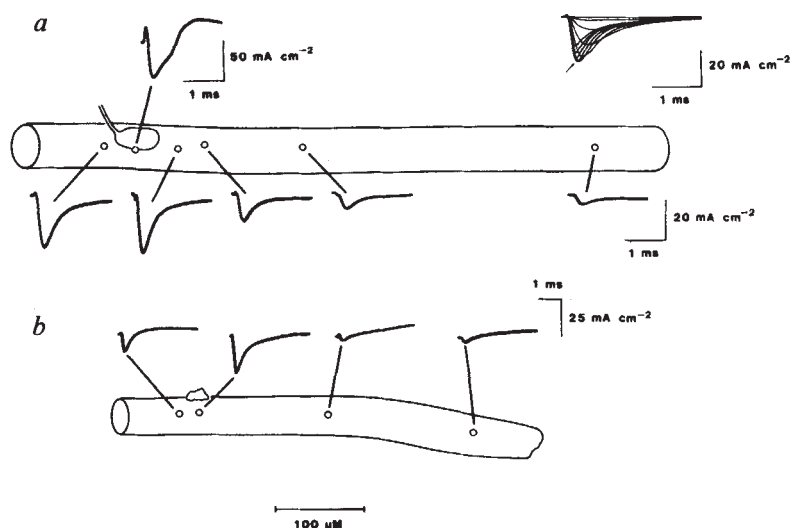
whereas in the fifth fibre it was 1.9 times larger. In four rat fibres, the ratio of peak Na current density at sites $\leq 30 \mu\text{m}$ from the endplate to Na current density at sites $\geq 140 \mu\text{m}$ away was 6.0 ± 2.1 (range 3.7–8.6, total of 7 near sites and 7 far sites).

The Na currents recorded at intermediate distances from the endplate are intermediate in size between the currents recorded near and far (Fig. 2); this is shown quantitatively in Fig. 3, where peak Na current amplitude is plotted as a function of distance from the endplate. For the snake fibre in Fig. 3*a*, the increased Na current density falls with a length constant of $\sim 80 \mu\text{m}$ and for another snake fibre (data not shown) the length constant was $\sim 90 \mu\text{m}$. Although the fragility of the isolated rat fibres prevented so detailed a mapping, the decline of Na current

‡ Present address: Mark O. Hatfield Marine Science Center, Oregon State University, Newport, Oregon 97365, USA.

Fig. 2 Representative Na currents recorded at different positions along *a*, a snake muscle fibre and *b*, a rat muscle fibre. Circles indicate the recording positions. The fibre maps are drawn to scale; the 100- μm calibration bar applies to *a* and *b*. *a*, Na currents, elicited by a test step to a potential estimated to be -22 mV , from six sites on a snake muscle fibre (number 0-6). Before measurement of the test currents at a spot, the membrane was held for 1-2 min at a potential estimated to be -120 mV . The pipette had a tip diameter of $10\text{ }\mu\text{m}$ and a resistance of $180\text{ k}\Omega$. The seal resistance varied from 350 to $590\text{ k}\Omega$. Note that the current calibration is different for the current illustrated above the map. This trace was recorded with the tip of the pipette partially overlapping the endplate. At this spot, steps of pipette potential of $>67.5\text{ mV}$ from the holding potential elicited not only a Na current but also caused an all-or-none twitch of the muscle fibre. The presence of a twitch implies that the large voltage-clamp pulse initiated a propagated action potential. It seems likely both that the action potential was the result of evoked transmitter release, as the pipette tip was lying partially over the nerve terminal, and that the action potential was responsible for the distortion of the current. The inset illustrates a superposition of test currents that were measured from the leftmost spot of the fibre map for test steps of 52.5 - 120 mV from the holding potential in 7.5-mV intervals; the arrow indicates the same peak test current illustrated below the leftmost spot of the fibre map. *b*, Na currents, elicited by a test pulse to a potential estimated to be -7 mV , from four sites on an isolated rat fibre (number D-15). The test currents were measured 2-4 min after the establishment of a holding potential estimated to be -97 mV . The pipette had a tip diameter of $5\text{ }\mu\text{m}$ and a resistance of $400\text{ k}\Omega$. The seal resistance varied from 600 - $900\text{ k}\Omega$.

Methods. Muscle fibres were dissected and bathed in a physiological solution containing (in mM with values given first for snake and then for rat): NaCl (145, 146), KCl (4.3, 5.0), CaCl_2 (3.5, 2), MgCl_2 (1.7, 1), glucose (11, 11), HEPES [10, titrated with NaOH to pH 7.0 (snake) or 7.4 (rat)]. Single fibres were dissociated enzymatically from the FDB muscle of 200-g male Sprague-Dawley rats by incubating the muscle for 1-2 h at 37°C in oxygenated physiological solution (described above) to which collagenase (2 mg ml^{-1} , Sigma, Type 1) and bovine serum albumin (1 mg ml^{-1} , Sigma, Fraction V) had been added. Single-barrelled pipettes were fabricated with two separate pulls on a vertical puller and fire-polished to a final tip diameter of 5 - $10\text{ }\mu\text{m}$. After good mechanical contact had been made between the pipette and the surface of the muscle fibre, the mechanical and electrical stability of the seal was improved further by applying suction of 10 - 20 (isolated rat fibres) or 100 - 120 mm Hg (snake muscle fibres). With minor modifications, the voltage-clamp circuit was that described in ref. 4, which provides analogue compensation for pipette and shunt resistances (R_{pipette} and R_{shunt}). An accurate measurement of pipette voltage and membrane current requires both compensation of the pipette resistance and an adequate shunt resistance. In control experiments, we found that R_{shunt} was adequate whenever $R_{\text{shunt}}/R_{\text{pipette}}$ exceeded ~ 1.2 , as the amplitude of the peak current measured from a patch was then independent of shunt resistance. Currents were amplified, filtered at 10 - 20 kHz by an 8-pole Bessel filter, digitized and stored on floppy diskettes of a DEC LAB 11/03 computer (Digital Equipment Corp.). The bulk of the linear leakage and capacity current was compensated electronically. Residual linear current was digitally compensated by subtracting an appropriately scaled control current from the test currents. Because the absolute transmembrane potential is not measured directly in the loose patch-clamp method, we estimated potential from the inactivation versus voltage relationship. Na current elicited by a constant test pulse was measured in response to 100 -ms prepulses (snake) or 40 -ms prepulses (rat) of varying amplitude. The prepulse potential inactivating half the Na current was assumed to occur at an absolute transmembrane potential of -75 mV . For accurate comparison of currents between patches, it is necessary to minimize differences in the level of inactivation from one patch to another. For this reason, we held at potentials from -120 to -100 mV (a single potential in a single fibre) for 1-2 min in the snake and 2-4 min in the rat. At the experimental temperatures used (20 - 22°C), these times and holding potentials produced sufficiently reproducible current amplitudes to allow comparison of currents from one spot to another.



density seemed to be steeper than for the snake fibres. Thus, for the fibre in Fig. 3*b* and one other fibre (data not shown), the length constant was $\sim 30\text{ }\mu\text{m}$. In addition to declining with longitudinal distance from the endplate, the peak Na current density also seemed to decrease with circumferential distance, although we were unable to examine this point in detail. In one isolated FDB fibre, from which the nerve terminal was dissociated completely, we were able to record directly from the region formerly covered by the terminal and found the peak Na current density to be ~ 35 times greater than at sites far from the endplate. This increased density may be accounted for in part by the junctional folds which increase the total membrane area of the endplate by a factor of ~ 5 (ref. 6).

In frog skeletal muscle, Almers *et al.*⁵ reported variations of as much as 3- to 4-fold in the amplitude of Na currents recorded from sites of unknown relationship to the endplate over distances as short as $20\text{ }\mu\text{m}$. Variations of about this magnitude are also apparent in the amplitude of currents measured far from the endplate (Fig. 3), but these spots were not spaced closely as in ref. 5. To explore the possibility that localized regions of high Na current density occur away from the endplate, we mapped current densities in five snake fibres at 4-6 spots spaced $10\text{ }\mu\text{m}$ apart and found up to a two-fold variation in peak current between spots. In no case, however, did the current density approach that measured near the endplate. Although we cannot exclude the possibility that we missed localized regions of greatly

increased Na current density at sites far from the endplate, we conclude that the important region of locally increased density is associated with the endplate.

In preliminary studies of the rat omohyoid muscle, we have measured current densities at sites far from the endplate comparable with those found by Almers *et al.*⁷ in the rat omohyoid. This omohyoid density is about fivefold higher than the density far from the endplate in the FDB. It will be of interest to determine whether the density of Na current near the endplate in the omohyoid is increased substantially over this already high distal Na current density. Another question of considerable interest is whether Na channels near the endplate have properties different from those far from the endplate. Na-dependent regenerative responses resistant to tetrodotoxin (TTX) and restricted to the endplate region have been observed in normally innervated frog⁸ and rat⁹ muscle, suggesting that there are TTX-resistant Na channels near the endplate. Our preliminary results with snake muscle give no indication of the presence of such TTX-resistant Na channels near the endplate. In addition, we have found no marked differences in the kinetics of Na current measured near and far from the endplate.

Our results are consistent with and explain two earlier observations: (1) intracellular recordings revealed an increased rate of rise of the action potential in the endplate region of frog¹⁰ and rat muscle^{9,11}; (2) extracellular current measurements with a vibrating microprobe^{12,13} revealed a steady inward current

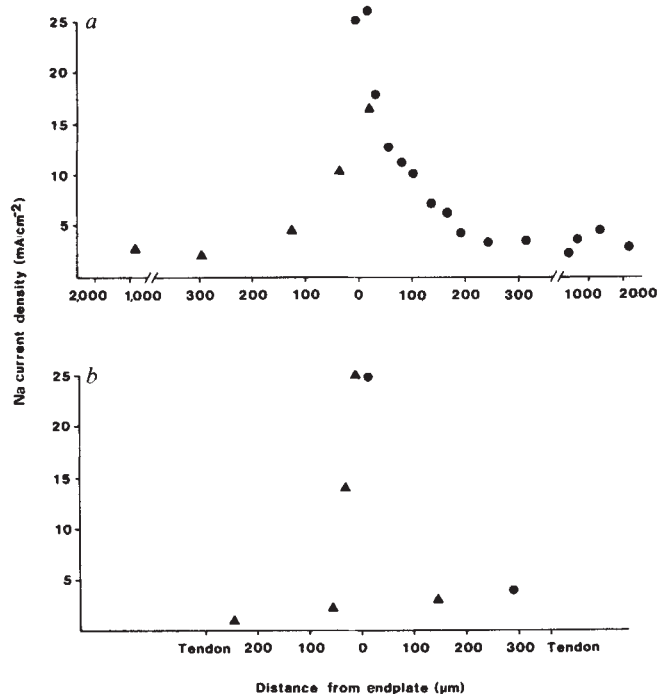


Fig. 3 Variation of sodium current density with distance from the endplate in *a*, a snake muscle fibre and *b*, a rat muscle fibre. Representative currents from these two muscle fibres are illustrated in Fig. 2 and the procedures for measuring the currents are described in Fig. 2 legend. The abscissa gives the distance from the recording site to the closest portion of the endplate. In *a*, note the change in horizontal scale for points at distances greater than 350 μm from the endplate. The points plotted as circles were measured sequentially at sites from left to right along the fibre and the points plotted as triangles were measured sequentially at sites from right to left along the fibre.

localized to the endplate region of veratridine-treated rat muscle¹¹. Both observations are consistent with our finding of an increased Na channel density near the endplate.

The density of several different kinds of ion channels seems to differ between regions near and far from the endplate. Thus, anomalous rectifier channels seem to occur at increased density near the endplate of frog fibres¹⁴ and Cl channels at reduced density near the endplate of rat fibres¹⁵, but there is insufficient information to determine whether this density alteration is more or less sharply localized to the endplate region than we describe here for the Na current. The distribution of ACh receptors has been mapped carefully in the snake² and rat¹ muscle preparations used here and, only 10 μm or less from the edge of the subsynaptic membrane, the sensitivity to ACh decreased by 15–50-fold compared with the subsynaptic membrane. The density of receptors outside the subsynaptic membrane undergoes a very slow, continuing decline. Thus, the distribution of ACh receptors differs from the distribution found by us for Na channels.

Our observation of an increased Na current density near the endplate region of rat and snake muscle raises important issues. For example, what is the functional significance of this increased density? One possibility is that by lowering the threshold for initiation of an action potential, the increased density raises the safety factor for impulse transmission. It will be important to understand the mechanisms responsible for creating and maintaining the elevated Na current density near the endplate. Clarification of the distribution of Na channels in developing muscle and in adult muscle following denervation should help to elucidate these mechanisms.

This research was supported by NIH grants NS 14901 (K.G.B.) and NS 16922 (J.H.C.) and grants from the Muscular Dystrophy Association of America to K.G.B. and D.T.C.

Received 6 August; accepted 20 December 1984.

1. Bekoff, A. & Betz, W. J. *J. Physiol. Lond.* **271**, 25–40 (1977).
2. Kuffler, S. W. & Yoshikami, D. *J. Physiol. Lond.* **244**, 703–730 (1975).
3. Stühmer, W. & Almers, W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 946–950 (1982).
4. Stühmer, W., Roberts, W. M. & Almers, W. in *Single-channel Recording* (eds Sakmann, B. & Neher, E. (Plenum, New York, 1983).
5. Almers, W., Stanfield, P. R. & Stühmer, W. *J. Physiol. Lond.* **336**, 261–284 (1983).
6. Salpeter, M. M. & Eldefrawi, M. E. *J. Histochem. Cytochem.* **21**, 769–778 (1973).
7. Almers, W., Roberts, W. M. & Ruff, R. L. *J. Physiol. Lond.* **347**, 751–768 (1984).
8. Nasledov, G. A., Volkov, E. M. & Poletaev, G. I. *Experientia* **38**, 576–577 (1982).
9. Thesleff, L., Vyskocil, F. & Ward, M. R. *Acta physiol. Scand.* **91**, 196–202 (1974).
10. Nastuk, W. L. & Alexander, J. T. *Fedn Proc.* **32**, 333 (1973).
11. Betz, W. J., Caldwell, J. H. & Kinnamon, S. C. *J. Physiol. Lond.* **352**, 189–202 (1984).
12. Jaffe, L. F. & Nuccitelli, R. *J. Cell Biol.* **63**, 614–628 (1974).
13. Betz, W. J. & Caldwell, J. H. *J. gen. Physiol.* **83**, 143–156 (1984).
14. Katz, B. & Miledi, R. *Proc. R. Soc. B216*, 497–507 (1982).
15. Betz, W. J., Caldwell, J. H. & Kinnamon, S. C. *J. gen. Physiol.* **83**, 175–192 (1984).

Autonomous chaotic behaviour of the slime mould *Dictyostelium discoideum* predicted by a model for cyclic AMP signalling

J. L. Martiel & A. Goldbeter

Faculté des Sciences, Université Libre de Bruxelles, Campus Plaine, CP 231, B-1050 Brussels, Belgium

How sustained oscillations lose their periodicity and thus give rise to chaos was first analysed in mathematical models^{1–3}, then observed in chemical systems such as the Belousov–Zhabotinsky reaction^{4,5} where chaos is autonomous because it originates from endogenous kinetic mechanisms. In contrast, chaos can also be obtained by periodically forcing an oscillatory system, as shown, for example, in cardiac cells⁶ and yeast glycolysis⁷. Biochemical evidence for autonomous chaos has been obtained both *in vitro* for the peroxidase reaction⁸ and in enzymatic models^{9,10} not based directly on experimental systems. We report here the occurrence of autonomous chaos in a realistic model for the cyclic AMP signalling system of the slime mould *Dictyostelium discoideum*^{11–13}, based on receptor modification. This model¹⁴ is also capable of bursting, a phenomenon characteristic of some pacemaker neurones such as R15 in *Aplysia*¹⁵. Whereas bursting has not been observed in *D. discoideum*, our model suggests that ‘aperiodic signalling’¹⁶ in the mutant *Fr17* provides the first example of autonomous chaos occurring spontaneously at the cellular level.

After starvation, *D. discoideum* amoebae aggregate around centres in a wave-like manner, via a chemotactic response to pulses of cyclic AMP emitted with a periodicity of several minutes by the centres and relayed towards the periphery of the aggregation field^{11–13,17–20}. (Periodic synthesis of cyclic AMP is also observed in cell suspensions²¹.) The molecular mechanism of cyclic AMP oscillations rests on the autocatalytic synthesis of cyclic AMP in *D. discoideum*^{22–24}; when cyclic AMP binds to a cell-surface receptor, adenylate cyclase is activated and the cyclic AMP synthesized is transported into the extracellular medium where it is hydrolysed by phosphodiesterase^{11–13}. The analysis of a model based on this regulation showed²⁴ that relay and oscillations can occur as a result of the positive feedback exerted by extracellular cyclic AMP and that both properties appear sequentially because of the continuous change in enzyme activities during development²⁵. In this model, ATP depletion acts as a counterpoise to autocatalysis throughout relay and oscillations. Experiments suggest, however, that significant ATP variation does not occur during cyclic AMP signalling²⁶ and that adaptation to constant cyclic AMP stimuli rather results from modification of a component in the pathway from receptors to adenylate cyclase²⁷. To account for these observations, we have analysed a model (Fig. 1) based on the common mechanism of receptor modification observed for various hormonal and chemotactic stimuli^{28,29}.

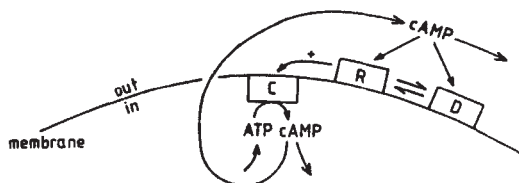


Fig. 1 Model based on receptor modification for the cyclic AMP signalling system of *Dictyostelium discoideum*. Binding of extracellular cyclic AMP to the active form of the receptor (R) elicits activation of adenylate cyclase (C). Arrows indicate synthesis of ATP, transport of the cyclic AMP into the extracellular medium and cyclic AMP hydrolysis by the intracellular and extracellular forms of phosphodiesterase. Decrease in cyclic AMP synthesis follows from the passage of the receptor into a desensitized state (D) uncoupled from adenylate cyclase¹⁴.

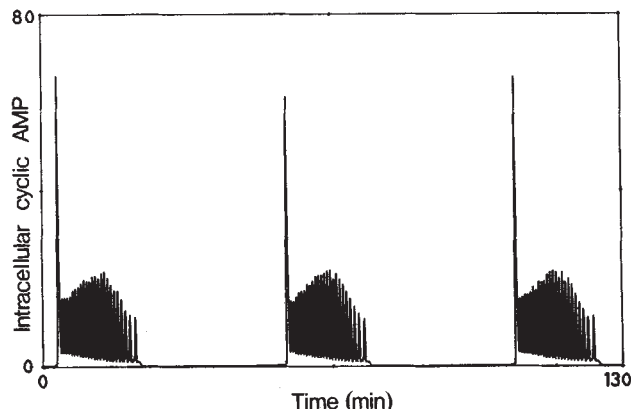


Fig. 2 Bursting. The time evolution of the normalized concentration of intracellular cyclic AMP is obtained by numerical integration of the seven kinetic equations (see ref. 14) which govern the model of Fig. 1. Parameter values are $\theta = \lambda = 0.01$, $\varepsilon = 0.2$, $\eta = \mu = 0.1$, $h = 5$, $L_1 = 300$, $L_2 = L_1/c^2$ with $c = 100$, $k_1 = 0.15 \text{ s}^{-1}$, $k_2 = 0.06 \text{ s}^{-1}$, $\sigma = 0.1 \text{ s}^{-1}$, $k_e = 0.8 \text{ s}^{-1}$, $k_i = 0.6 \text{ s}^{-1}$, $k_t = 0.4 \text{ s}^{-1}$, $d_1 = d_2 = d_3 = 10 \text{ s}^{-1}$, $v = 7.5 \times 10^{-3} \text{ s}^{-1}$, $q = 4,000$ and $K_R^{1/2} = 10^{-7} \text{ M}$.

In the model shown in Fig. 1, the increase in cyclic AMP in the course of relay and oscillations still results from receptor-mediated activation of adenylate cyclase by extracellular cyclic AMP, but the decreasing phase is brought about mainly by the passage of the receptor from an active to a desensitized state uncoupled from adenylate cyclase. The latter state represents either another conformation, as for the acetylcholine receptor^{30,31}, or a form modified covalently, for example, by reversible phosphorylation³². The model is governed by seven kinetic equations (see ref. 14 for these equations and a definition of the parameters). When the steady state admitted by these equations becomes unstable, the cyclic AMP signalling system generally operates periodically. The oscillations of intracellular and extracellular cyclic AMP occur in the absence of significant ATP variation and are accompanied by a periodic shift of the receptor from the active to the desensitized state¹⁴. Such behaviour can be associated with the periodic release of cyclic AMP pulses by aggregation centres. Relay of cyclic AMP pulses is observed for proximate parameter values. The model also accounts¹⁴ for adaptation of the signalling system to constant cyclic AMP stimuli³³.

Complex oscillatory behaviour also occurs in this mechanism, first in the form of bursting (Fig. 2) in which series of sharp cyclic AMP pulses periodically alternate with phases of quiescent behaviour. For slightly different values of the parameters, the system admits two simultaneously stable periodic regimes; similar birhythmicity^{9,10} has been observed recently in two chemical systems^{5,34}. It is in a proximate parameter domain that the system undergoes a transition to chaos. The oscillations become highly irregular (Fig. 3) while the system evolves towards a 'strange' attractor in the phase space (Fig. 4). On such a

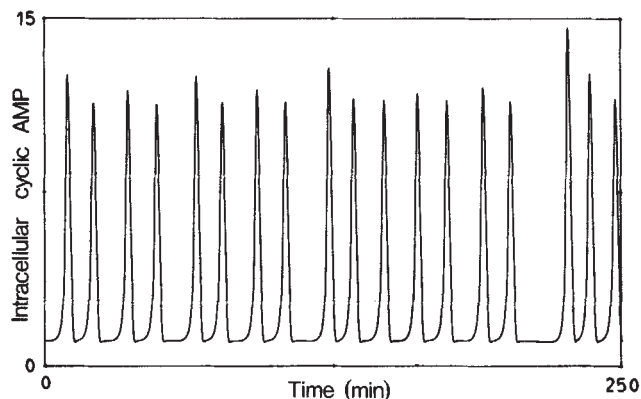


Fig. 3 Aperiodic oscillations (chaos) in cyclic AMP synthesis. The average time interval between successive peaks compares with the value near 14 min found¹⁶ for aperiodic signalling in *D. discoideum* mutant *Fr17*. Only slight variations in ATP accompany the chaotic synthesis of cyclic AMP (see Fig. 4). The curve is obtained as in Fig. 2 for $v = 7.545 \times 10^{-4} \text{ s}^{-1}$; other parameter values are as in Fig. 2 but the constants in s^{-1} are divided by 10.

trajectory the system never passes twice through the same point and is highly sensitive to initial conditions, because two points initially close to each other will exponentially diverge in time. As shown by the histogram in Fig. 4, the length of one cycle on the attractor varies by a factor of up to three in an apparently random manner (Fig. 3).

The origin of bursting and chaos may be understood in molecular terms by the fact that the model contains two oscillatory mechanisms sharing a unique feedback loop. The two mechanisms differ by the process limiting the autocatalysis of cyclic AMP, this limitation resulting alternatively from substrate depletion and desensitization as the receptor progressively passes into the state uncoupled from adenylate cyclase on incubation with extracellular cyclic AMP. When either one of these effects predominates, the system shows simple periodic oscillations; complex oscillatory phenomena appear when the two effects acquire comparable importance. As for regular oscillations¹⁴, however, the aperiodic oscillatory synthesis of cyclic AMP is accompanied by only slight variations in ATP (see Fig. 4). In a previous model, birhythmicity, chaos and complex periodic oscillations resulted from the coupling in series of two autocatalytic enzyme reactions^{9,10}. Here, these phenomena arise when two oscillatory mechanisms sharing the same feedback loop are coupled in parallel. In this latter model, receptor modification has a primary role in the mechanism of oscillatory phenomena. As well as their application to cellular slime moulds, the present results could relate to other processes where oscillatory phenomena and receptor desensitization occur, for example, in hormonal regulation and neural function.

Bursting has not yet been observed in *D. discoideum* amoebae, but some observations on a mutant of *D. discoideum* may be interpreted in terms of chaos by our present analysis. Durston¹⁶ has measured the time intervals between successive waves in the wild type and in two pacemaker mutants aggregating on agar. In contrast to the wild type, the mutant *Fr17* is characterized by irregular time intervals between waves, suggesting that *Fr17* centres have the property of aperiodic signalling¹⁶. The average as well as the range of variation of the time interval compare with those obtained in Figs 3 and 4. A preliminary study of *Fr17* in cell suspensions showed the occurrence of 'erratic' oscillations of cyclic AMP³⁵, but the reported time course extended over a few cycles only. The results described in Figs 3 and 4 suggest that the aperiodic behaviour of *Fr17* during aggregation on agar or in cell suspensions may represent chaos. The mutation in *Fr17* affects cyclic AMP metabolism³⁶ and is associated with a precocious activation of adenylate cyclase after starvation³⁵. Thus, the parameters governing cyclic AMP synthesis could be such that the dynamic behaviour of

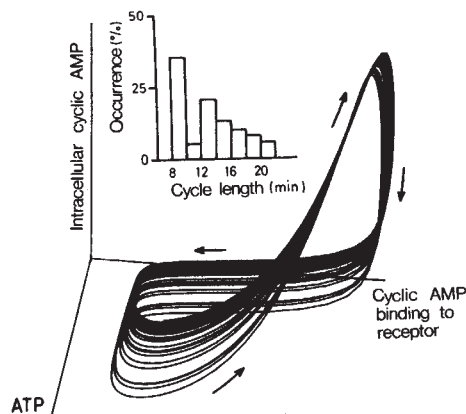


Fig. 4 Strange attractor in the model for the cyclic AMP signalling system based on receptor modification. Shown is the trajectory followed by the system in the conditions of Fig. 3, projected in the phase space formed by the normalized concentrations of intracellular cyclic AMP (range 0–15), ATP (1.545–1.585) and the fraction of receptor in the R and D states binding cyclic AMP (0.15–0.9). The insert shows a histogram of cycle lengths for the strange attractor; the time between two successive peaks of intracellular cyclic AMP was measured for ~500 cycles and the percentage of cycle lengths in 2-min intervals plotted.

the signalling system is chaotic in the mutant *Fr17* whereas it is periodic in the wild type.

This hypothesis could be tested by a more detailed study of cyclic AMP synthesis in *Fr17* suspensions and, if confirmed, would provide the first example of autonomous chaos at the cellular level in physiological conditions. The *D. discoideum* mutant would also exhibit the equivalent, for a cellular system, of a 'dynamical disease'³⁷ in which abnormal modes of dynamic behaviour arise in an intact physiological system from a variation in parameter values. Whereas such variation originates from mutation in *Fr17*, the transition to chaos has been induced in a molluscan neurone by means of drugs³⁸.

Changes in enzyme activities may cause the developmental transitions no relay-relay-oscillations observed in the *D. discoideum* wild type after starvation²⁵. Such changes may also explain why the signalling system of *Fr17* sometimes leaves the aperiodic domain and enters a more regular regime during development¹⁶. In agreement with the former suggestion²⁵, the addition of exogenous phosphodiesterase initiates periodic aggregation in a *D. discoideum* mutant lacking this enzyme³⁹. The present model suggests that the addition of phosphodiesterase to suspensions of oscillating *Fr17* cells or *Fr17* amoebae aggregating on agar should induce a transition from chaotic to periodic behaviour in an analogous manner.

This work was carried out while J.L.M. held a European Community Fellowship.

Received 17 September; accepted 21 November 1984.

1. Lorenz, E. N. *J. atmos. Sci.* **20**, 130–141 (1963).
2. May, R. *Nature* **261**, 459–467 (1976).
3. Feigenbaum, M. J. *J. statist. Phys.* **19**, 25–52 (1978).
4. Schmitz, R. A., Graziani, K. R. & Hudson, J. L. *J. chem. Phys.* **67**, 3040–3044 (1977).
5. Roux, J. C. *Physica* **7D**, 57–68 (1983).
6. Guevara, M. R., Glass, L. & Shrier, A. *Science* **214**, 1350–1353 (1981).
7. Markus, M., Kuschmitz, D. & Hess, B. *FEBS Lett.* **172**, 235–238 (1984).
8. Olsen, L. F. & Degn, H. *Nature* **267**, 177–178 (1977).
9. Decroly, O. & Goldbeter, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6917–6921 (1982).
10. Goldbeter, A. & Decroly, O. *Am. J. Physiol.* **245**, R478–R483 (1983).
11. Gerisch, G. & Malchow, D. *Adv. cyclic Nucleotide Res.* **7**, 49–68 (1976).
12. Newell, P. C. in *Microbial Interactions: Receptors and Recognition* Ser B, Vol. 3 (ed. Reissig, J. L.) 3–57 (Chapman & Hall, London, 1977).
13. Gerisch, G. *Rev. Physiol.* **44**, 535–552 (1982).
14. Martiel, J. L. & Goldbeter, A. *C. r. hebdo. Séanc. Acad. Sci., Paris, Ser. III* **298**, 549–552 (1984).
15. Alving, B. O. *J. gen. Physiol.* **51**, 29–45 (1968).
16. Durston, A. *J. Dev. Biol.* **38**, 308–319 (1974).
17. Shaffer, B. M. *Adv. Morphogen.* **2**, 109–182 (1962).
18. Konijn, T. M., Van de Meene, J. G. C., Bonner, J. T. & Barkley, D. S. *Proc. natn. Acad. Sci. U.S.A.* **58**, 1152–1154 (1967).
19. Alcantara, F. & Monk, M. *J. gen. Microbiol.* **85**, 321–334 (1974).
20. Tomchik, K. J. & Devreotes, P. N. *Science* **212**, 443–446 (1981).
21. Gerisch, G. & Wick, U. *Biochem. biophys. Res. Commun.* **65**, 364–370 (1975).
22. Goldbeter, A. *Nature* **253**, 540–542 (1975).

23. Goldbeter, A. & Caplan, S. R. *A. Rev. Biophys. Bioengng* **5**, 449–476 (1976).
24. Goldbeter, A. & Segel, L. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1543–1547 (1977).
25. Goldbeter, A. & Segel, L. A. *Differentiation* **17**, 127–135 (1980).
26. Roos, W., Scheidegger, C. & Gerisch, G. *Nature* **266**, 259–261 (1977).
27. Theibert, A. & Devreotes, P. N. *J. Cell Biol.* **97**, 173–177 (1983).
28. Stadel, J. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 3173–3177 (1983).
29. Springer, M. S., Goy, M. F. & Adler, J. *Nature* **280**, 279–284 (1979).
30. Katz, B. & Thesleff, S. *J. Physiol., Lond.* **138**, 63–80 (1957).
31. Heidmann, T. & Changeux, J. P. *A. Rev. Biochem.* **47**, 317–357 (1978).
32. Lubs-Haukeness, J. & Klein, C. *J. biol. Chem.* **257**, 12204–12208 (1982).
33. Devreotes, P. N. & Steck, T. L. *J. Cell Biol.* **80**, 300–309 (1979).
34. Alamgir, M. & Epstein, I. R. *J. Am. chem. Soc.* **105**, 2500–2501 (1983).
35. Coukell, M. B. & Chan, F. K. *FEBS Lett.* **110**, 39–42 (1980).
36. Kessin, R. H. *Cell* **10**, 703–708 (1977).
37. Mackey, M. C. & Glass, L. *Science* **197**, 287–289 (1977).
38. Holden, A. V., Winlow, W. & Haydon, R. G. *Biol. Cybern.* **43**, 169–173 (1982).
39. Darmon, M., Barra, J. & Brachet, P. *J. Cell Sci.* **31**, 233–243 (1978).

Ontogeny of the T-cell antigen receptor within the thymus

H. Ralph Snodgrass, Pawel Kisielow, Michael Kiefer, Michael Steinmetz & Harald von Boehmer

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland

The expression of T-cell antigen receptors during T-cell ontogeny is an important issue that bears directly on such questions as where T-cell tolerance is acquired, at what stage T cells become susceptible to repertoire selection, and why most thymocytes die within the thymus. The thymus rudiment is colonized during days 11 and 12 of gestation¹, but it is not until day 19 that significant numbers of functional thymocytes are present^{2,3}. Although much is known about the ontogeny of function- and specificity-associated surface molecules such as Ly2 and MT4 (the murine equivalent of human T4) during this period^{3,4}, the ontogeny of the T-cell antigen receptors remains obscure. We have now addressed this question on three levels: DNA rearrangement, messenger RNA transcription and expression of cell-surface receptor-like proteins. Our results suggest that T-cell receptors are first expressed within the thymus around day 17 of gestation, independently of and probably before the expression of Ly2 and MT4. Furthermore, these data suggest that all major adult thymocyte subpopulations, including the small cortical cells, most of which die within the thymus^{5,6}, express receptors.

The genes coding for one chain of the T-cell receptor have recently been isolated^{7–9}. These genes are rearranged and expressed only in T cells and are related to immunoglobulin genes by sequence and genomic organization^{10,11}. They are known to encode the β -chain of the T-cell receptor through the identity between the N-terminal amino acid sequence and that predicted from complementary DNA (cDNA) clones^{9,12}. As with immunoglobulins^{13,14}, genetic rearrangements are a prerequisite for expression of a functional T-cell receptor^{15,16}. The extent of rearrangement of the β -chain gene complex can most easily be evaluated using a cDNA constant-region (C) probe and *Hind*III-digested DNA. Such a combination will detect a 9.5- and a 3.0-kilobase (kb) band containing $C_{\beta}1$ and $C_{\beta}2$, respectively (Fig. 1c). Assuming a deletion model, rearrangements to either locus will result in a depletion of the 9.5-kb band (either by deletion or by altering the fragment size), whereas the 3-kb band will be unchanged and serves as an internal control. The change in ratio between these two bands is an indication of the extent of rearrangement of the β locus. Analysis of DNA from germline cells showed the two bands to be of equal intensity, whereas with DNA from a T lymphoma, BW5147, in which the $C_{\beta}1$ locus has been deleted on both chromosomes, only the 3-kb band was seen (Fig. 1a). Thymocytes at day 15 of gestation showed some evidence of rearrangements, indicated not only by the depletion of the 9.5-kb band, but also by the appearance of rearranged bands predominantly of lower relative molecular mass (M_r). The intensity of the 9.5-kb band fell abruptly at day 17 of gestation, and in

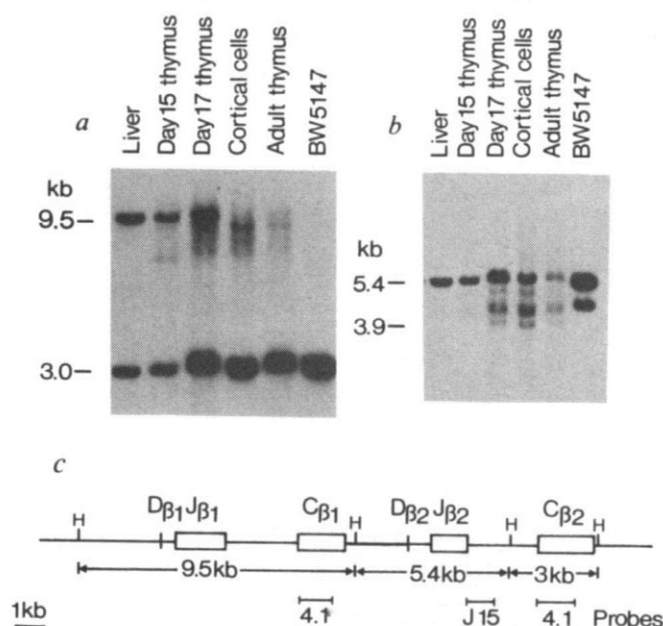


Fig. 1 Southern blot analysis of the T-cell receptor β -chain genes in cells from embryonic and adult mouse thymus. *a*, DNAs digested with *Hind*III were separated on a 0.6% agarose gel and transferred to a Zeta probe membrane (Bio-Rad). Hybridization was with a nick-translated³², gel-purified insert of the 4.1 cDNA clone, containing ~70 bp of a V-region gene segment, a J-gene segment and ~400 bp of either the $C_{\beta 1}$ or the $C_{\beta 2}$ gene of the T-cell receptor β -chain. This cDNA clone was isolated with a 33-residue oligonucleotide, specific for the C_{β} gene segment, from a cDNA library of the KdD4 cytotoxic T-cell line (M.K. and M.S., unpublished results). *b*, After removal of the 4.1 probe, the same filter was hybridized with gel-purified insert of the J15 subclone. This fragment was obtained from a λ phage genomic clone (CSp4) isolated from a cytotoxic T-cell line (CSp2) by double digestion with *Cl*aI and *Eco*RI and subcloned into pUC9 (H.R.S., unpublished results). It contains the 900-bp *Cl*aI/*Eco*RI fragment located ~90 bp downstream of the $J_{\beta 2}$ gene locus. *c*, Partial restriction map for the β -chain gene locus displaying the two D-gene segments ($D_{\beta 1}$ and $D_{\beta 2}$), the two clusters of J-gene segments ($J_{\beta 1}$ and $J_{\beta 2}$) and the two constant-region gene segments ($C_{\beta 1}$ and $C_{\beta 2}$). The relevant *Hind*III sites (H) are indicated together with the lengths of the resulting fragments and the regions hybridizing to probes 4.1 and J15 (refs 10, 11 and M.S., unpublished results).

Methods. Fetal thymocytes were obtained from timed matings where the day of vaginal plug detection was designated as day zero. Cortical cells were isolated from adult thymocytes by treatment with peanut agglutinin and anti-H-2K (100-5/28)³³ plus complement³⁴. The resulting cells were >98% small Ly2⁺MT4⁺ thymocytes (data not shown). The T-cell lymphoma BW5147 was maintained in tissue culture. DNAs were isolated from cell suspensions or fresh tissue according to Blin and Stafford³⁵ and digested to completion with *Hind*III. Probes were labelled by nick-translation (4.1) or the oligo-labelling method (J15)³⁶ to high specific activities (10^8 – 10^9 c.p.m. per μ g). Hybridization to ~5 μ g of DNA per lane was in 50% formamide, 1 M NaCl, 10% dextran sulphate, 0.2% bovine serum albumin (BSA), 0.2% Ficoll, 0.2% polyvinylpyrrolidone, 50 mM Tris-HCl (pH 7.4), 0.1% Na₂P₂O₇, 0.1% SDS, 150 μ g ml⁻¹ sheared and denatured salmon sperm DNA with 3×10^6 c.p.m. ml⁻¹ hybridization solution for 16 h at 42°C. After hybridization, filters were washed at 68°C twice in $3 \times$ SSC, 0.1% SDS and twice in $0.3 \times$ SSC, 0.1% SDS. Each wash was for 15 min. Exposure was overnight with two intensifying screens (Ilford fast tungstate). To remove the probe, the filter was incubated twice for 15 min at 90°C in $0.1 \times$ SSC, 0.1% SDS. Sizes of restriction fragments hybridizing to the probes were measured using phage λ DNA restriction fragments of known length run in parallel on the same gel.

the adult, essentially no germline band remained in either the whole thymocyte population or the small cortical Ly2⁺ MT4⁺ subpopulation (Fig. 1*a*). Densitometric analysis, in the linear range of film responses, indicated that in ~30% of day 15, 75% of day 17 and >95% of adult thymocytes, the β -chain locus had become rearranged (data not shown).

This type of analysis gives no indication of whether the rearrangements have occurred to $C_{\beta 1}$ or $C_{\beta 2}$. Such information is derived from *Hind*III-digested DNA analysed with a probe isolated from the joining region ($J_{\beta 2}$)– $C_{\beta 2}$ intron. This combination detects a 5.4-kb germline fragment that will be altered by $J_{\beta 2}$ rearrangements (Fig. 1*c*). Rearrangements of $J_{\beta 2}$ were first apparent at day 17, and in the adult essentially all thymocytes had $J_{\beta 2}$ rearrangements (Fig. 1*b*). It is not known what percentage of these are diversity (D)– $J_{\beta 2}$ rearrangements and what percentage are variable (V)– D – $J_{\beta 2}$ rearrangements.

Northern blot analysis of the β -chain locus during thymus ontogeny showed that the transcriptional activity paralleled the extent of DNA rearrangement during ontogeny. There was relatively little mRNA at day 15, but this increased progressively with time, reaching a maximum in the neonate and adult thymus. In contrast, levels of mRNA in peripheral T cells fall dramatically (Fig. 2*a, c*). This difference levels of β -chain message between thymocytes and peripheral T cells has been observed by others¹⁷. The role of such a high level of mRNA during thymocyte differentiation is unknown and seems to be the opposite of the situation observed in B-cell development, where levels of immunoglobulin mRNA peak during the terminal plasma cell stage¹⁸. Small cortical thymocytes contain only one-sixth of the RNA present in the whole thymocyte population (data not shown), but the β -chain mRNA represents the same fraction of the total RNA in both populations (Fig. 2*c*). Analysis by fluorescence-activated cell sorting revealed that the subpopulation of small cortical thymocytes was contaminated with <2% of non-cortical cells (data not shown). Although the mRNA detected in this subpopulation may have originated from a few contaminating cells expressing very high amounts of β -chain RNA, we consider this unlikely because our data at the protein level (Fig. 3*c*) suggest that a substantial fraction of cortical thymocytes express receptors. This conclusion is also supported by the observation that two thymomas with cortical phenotypes studied showed rearranged β -chain genes and expressed receptor-like protein (data not shown).

More revealing in terms of the expression of the T-cell receptor is the change in the ratio between the 1.3- and the 1.0-kb mRNA which occurred during ontogeny (Fig. 2). The 1.3-kb RNA represents a full-length transcript of the V–D–J–C elements, whereas the 1.0-kb message represents a transcript of only D–J–C elements^{19,20}. At day 15, most of the transcription was of non-functional D–J–C type, whereas at day 17 and in adult thymus, most of the mRNA was of the full-length V–D–J–C type.

T-cell antigen-specific receptors are disulphide-bonded heterodimeric molecules with subunits of M_r 40–50,000 (refs 21–25). Such molecules can be identified by diagonal gel electrophoresis^{26,27}. Several arguments support the assumption that molecules analysed by this technique are T-cell surface-antigen receptors. First, such molecules are T-cell-specific (ref. 27 and H.R.S., unpublished observations). Second, they have the same physical characteristic as molecules currently accepted to be the T-cell antigen receptor. Third, the only other reported molecule possessing similar physical properties is the Ly2 complex²⁸, but this molecule is easily distinguished from receptor-like molecules (Fig. 3*b*). Furthermore, the contribution of Ly2 to receptor-like spots was excluded by preclearing with monoclonal anti-Ly2 (data not shown).

The expression of cell-surface receptor-like molecules paralleled DNA rearrangements and transcription levels. At day 15 embryonic thymocytes did not have detectable amounts of receptor-like molecules, whereas by day 17 such molecules were readily demonstrable (Fig. 3*a, b*). The appearance of these molecules does not correlate with the expression of Ly2 or MT4,

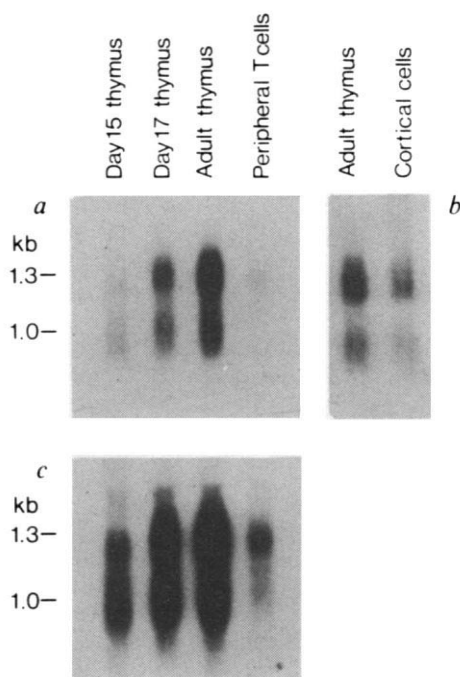


Fig. 2 Northern blot analysis of T cells at various stages of development. Total RNA (*a*, 20 μ g; *b*, 10 μ g) was separated on 1.1% glyoxal gels³⁷, transferred to Gene Screen (NEN) and hybridized with the 4.1 probe (Fig. 1*c*). *c* Represents a 12-fold longer exposure of *a*.

Methods. Embryonic thymocytes were obtained as described in Fig. 1 legend. Peripheral T cells were cells purified by passage through nylon wool. Cortical cells were isolated by treatment of adult thymocyte with monoclonal anti-H-2K (100-5/28) and passed through a protein A-Sepharose 6MB column (Pharmacia) at 4°C. More than 98% of cells in the non-adherent fraction were small Ly2⁺MT4⁺ thymocytes (data not shown). RNA was isolated by the guanidinium isothiocyanate and CsCl gradient centrifugation procedures^{38,39}. RNA blots were hybridized in 1 M NaCl, 50 mM Tris pH 7.5, 1% SDS, 0.2% polyvinylpyrrolidone, 0.2% BSA, 0.2% Ficoll, 100 μ g ml⁻¹ denatured salmon sperm DNA, 10% dextran sulphate and 1–5 $\times 10^6$ c.p.m. ml⁻¹ oligo-labelled³³ 4.1 probe (10⁹ c.p.m. per μ g). Filters were washed in 0.3 M NaCl, 60 mM Tris pH 8, 2 mM EDTA, twice at 25°C for 10 min and twice at 68°C for 30 min and twice in 10 mM NaCl, 6 mM Tris pH 8, 2 mM EDTA at 25°C. Sizes of RNA molecules were estimated from denatured phage DNA restriction fragments of known lengths run in parallel on the same gel.

because at day 17 both the Ly2⁺MT4⁺ and the Ly2⁺MT4⁺ subpopulations expressed receptor-like molecules, as deduced from the observation that, on a per cell basis, there was no significant difference between the amount of these molecules in the unfractionated and the Ly2⁺MT4⁺ subpopulation (Fig. 3*b, c*). Furthermore, in a panel of thymomas, β -chain DNA rearrangements and receptor-like molecules were observed in all four groups defined by the Ly2MT4 phenotype, including both the Ly2⁺MT4⁺ and Ly2⁺MT4⁺ subpopulations (H.R.S. and M.S., unpublished data).

In the adult thymus, receptor-like molecules were found on both the cortisone-resistant thymocytes and the small Ly2⁺MT4⁺ cortical thymocytes (Fig. 3*d, e*). Although it is very difficult to quantify this type of analysis, the assay is very sensitive (Fig. 3 legend), and we estimate roughly that >20% of the small cortical Ly2⁺MT4⁺ thymocytes express receptors.

Our conclusion concerning cell-surface receptor expression is consistent with data reported by Roehm *et al.*²⁹ but not with the conclusions of Acuto *et al.*³⁰ that T-cell receptors are expressed very late during T-cell ontogeny, concomitantly with the division into functional T8⁺T4⁺ and T8⁺T4⁺ subpopulations. These data were based on the correlation between the presence of the receptor and T3. We cannot test this question directly, however, because no murine equivalent of T3 has yet been

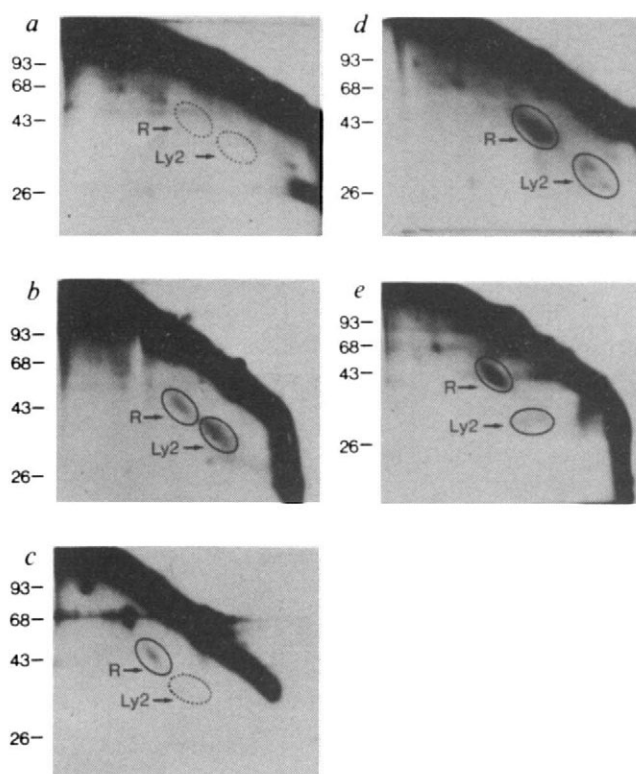


Fig. 3 Diagonal gel electrophoresis of thymocyte cell-surface proteins from day 15 (*a*), day 17 (*b*) and day 17 Ly2⁺MT4⁺ embryonic thymocytes (*c*), cortisone-resistant adult thymocytes (*d*) and small Ly2⁺MT4⁺ cortical adult thymocytes (*e*). The first dimension is non-reducing (left to right) and the second dimension is reducing (top to bottom). The presence of receptor-like proteins (R) or Ly2 is indicated by a solid line, whereas the absence of these proteins is indicated by a dotted line.

Methods. Embryonic and adult thymocytes were isolated as described in Fig. 1 legend. The Ly2⁺MT4⁺ subpopulation was isolated from day 17 embryonic thymocytes by a combined treatment with monoclonal anti-Ly2.2 (H02.2)³³ and anti-MT4 (M129.19)⁴⁰ plus complement. The M129.19 antibody was rendered cytotoxic by haptentation with trinitrophenol (TNP) and incubation with monoclonal anti-TNP IgM. The resulting cells were >98% large Ly2⁺MT4⁺ thymocytes (data not shown). Cortisone-resistant thymocytes were obtained 48 h after intraperitoneal injections of 2.5 mg cortisone acetate. Cortical thymocytes were isolated as described in Fig. 1 legend. Following cell-surface lactoperoxidase-catalysed iodination⁴¹, the cells were lysed (0.5% NP40, 150 mM NaCl, 10 mM Tris pH 7.2, 0.2% Na₂S₂O₃, 1 mM EDTA, 1 mM phenylmethylsulphonyl fluoride, 100 mM iodoacetamide) and the nuclei removed. Lysates from 4 $\times 10^6$ cells were analysed by diagonal gel electrophoresis^{26,27}. The first and second dimensions were non-reducing 7.5% and reducing 10% Laemmli gels⁴², respectively. The gels were dried and autoradiographed with Kodak XRP film and Ilford fast tungstate screens. Prestained *M_r* markers (BRL) were run in parallel on each gel. The sensitivity of this assay was estimated by mixing experiments with receptor-positive cells (cortisone-resistant thymocytes) and a receptor-negative thymoma (BW5147). The receptor-positive population had to be present at a frequency of >20% to be detected (data not shown).

described, but if the correlation between T3 and receptor expression holds, we predict that in the mouse the frequency of T3-positive thymocytes would be much higher than initially reported for humans.

Our data suggest that at day 15 of gestation thymocytes do not express T-cell receptors, that by day 17 a significant fraction of thymocytes express receptors, and that in the adult most thymocytes are expressing receptors. The fact that day 15 thymocytes differentiate, express receptors and become functional cells in organ cultures³¹ suggests that the rapid transition from receptor-negative to receptor-positive occurs within the thymus and is not a result of the entrance of receptor-positive cells.

We thank Carina Olsson, Katrin Schubiger, Verena Stauffer and Gholam Dastoornikoo for technical assistance; Wilhelm Bannwarth for synthesizing the oligonucleotide probe used in this study; Carolyn Hägi for preparing this manuscript; and Werner Haas for helpful discussions throughout this study. P.K. is on leave of absence from the Ludwik Hirsfeld Institute of Immunology and Experimental Therapy of the Polish Academy of Sciences, Wrocław, Poland. The Basel Institute for Immunology is supported by F. Hoffmann-La Roche and Co., CH-4005 Basel, Switzerland.

Received 2 November 1984; accepted 8 January 1985.

- Moore, M. & Owen, J. *J. exp. Med.* **126**, 715-725 (1967).
- Widmer, M., MacDonald, H. & Cerottini, J. *Thymus* **2**, 245-255 (1981).
- Ceredig, R., Dialynas, D., Fitch, F. & MacDonald, H. *J. exp. Med.* **158**, 1654-1671 (1983).
- Ceredig, R., MacDonald, H. & Jenkinson, E. *Eur. J. Immun.* **13**, 185-190 (1983).
- Cantor, H. & Weissman, I. *Prog. Allergy* **20**, 1-64 (1976).
- McPhee, D., Pye, J. & Shortman, K. *Thymus* **1**, 151-155 (1979).
- Hedrick, S., Cohen, D., Nielson, E. & Davis, M. *Nature* **308**, 149-153 (1984).
- Hedrick, S., Nielson, E., Kavalier, J., Cohen, D. & Davis, M. *Nature* **308**, 153-158 (1984).
- Yanagi, Y. *et al. Nature* **308**, 145-149 (1984).
- Gascoigne, N., Chien, Y., Becker, D., Kavalier, J. & Davis, M. *Nature* **310**, 387-391 (1984).
- Malissen, M. *et al. Cell* **37**, 1101-1110 (1984).
- Acuto, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3851-3855 (1984).
- Honjo, T. *A. Rev. Immun.* **1**, 499-528 (1983).
- Tonegawa, S. *Nature* **302**, 575-581 (1983).
- Chien, Y., Gascoigne, R., Kavalier, J., Lee, N. & Davis, M. *Nature* **309**, 322-326 (1984).
- Siu, J. *et al. Cell* **37**, 393-401 (1984).
- Yoshikai, Y., Yanagi, Y., Suciu-Foca, N. & Mak, T. *Nature* **310**, 506-508 (1984).
- Perry, R. & Kelley, D. *Cell* **18**, 1333-1339 (1979).
- Clark, S. *et al. Nature* **311**, 387-389 (1984).
- Siu, J. *et al. Nature* **311**, 344-350 (1984).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Samelson, L., Germain, R. & Schwartz, R. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6972-6976 (1983).
- Kappler, J., Kubo, R., Haskins, K., White, J. & Marrack, P. *Cell* **34**, 727-737 (1983).
- McIntyre, B. & Allison, J. *Cell* **34**, 739-746 (1983).
- Meuer, S. *et al. J. exp. Med.* **157**, 705-719 (1983).
- Hynes, R. & Destree, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2855-2859 (1977).
- Goding, J. & Harris, A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4530-4534 (1981).
- Ledbetter, J., Seaman, W., Tsu, T. & Herzenberg, L. *J. exp. Med.* **153**, 1503-1516 (1981).
- Roehm, N. *et al. Cell* **38**, 577-584 (1984).
- Acuto, O. *et al. Cell* **34**, 717-726 (1983).
- Kisielow, P., Leiserson, W. & von Boehmer, H. *J. Immun.* **133**, 1117-1123 (1984).
- Rigby, P., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237-251 (1977).
- Gottlieb, P. *et al. Immunogen* **10**, 545-555 (1980).
- Kisielow, P., von Boehmer, H. & Haas, W. *Eur. J. Immun.* **12**, 463-467 (1982).
- Blin, N. & Stafford, D. *Nucleic Acids Res.* **3**, 2303-2308 (1976).
- Feinberg, A. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
- McMaster, J. & Carmichael, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835-4838 (1977).
- Chirgwin, J., Przybyla, A., MacDonald, R. & Rutter, W. *Biochemistry* **18**, 5294-5299 (1979).
- Glisin, V., Crkvenjakov, R. & Byus, C. *Biochemistry* **13**, 2633-2637 (1977).
- Pierres, A. *et al. J. Immun.* **132**, 2775-2782 (1984).
- Vitetta, E. & Uhr, J. *Biochim. biophys. Acta* **415**, 253-271 (1975).
- Laemmli, U. *Nature* **227**, 680-685 (1971).

Expression of a minichromosomal variant surface glycoprotein gene in *Trypanosoma brucei*

Victoria Rothwell, Robert Aline Jr, Marilyn Parsons*, Nina Agabian* & Kenneth Stuart

Issaquah Health Research Institute, 1595 NW Gilman Boulevard, Issaquah, Washington 98027, USA

* Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

African trypanosomes contain numerous variant surface glycoprotein (VSG) genes, but express only one at a time¹. When different VSG genes are activated by gene duplicative²⁻⁴ or non-duplicative mechanisms⁵, antigenic variation occurs. Although transcriptionally inactive VSG genes can have either an intrachromosomal or a telomeric location⁶, all expressed VSG genes so far examined are telomeric. Some VSG genes are located on minichromosomes of ~100 kilobases (kb)⁷⁻⁹, but all those thus far described are transcriptionally inactive. We have found that the single copy of the IsTat 1.1 VSG gene in the IsTaR 1 serodeme has a telomeric location on a minichromosome when inactive. When it is activated, it is not duplicated and remains on a minichromosome. These data indicate that minichromosomes contain or can acquire all the DNA sequences necessary in *cis* for VSG gene

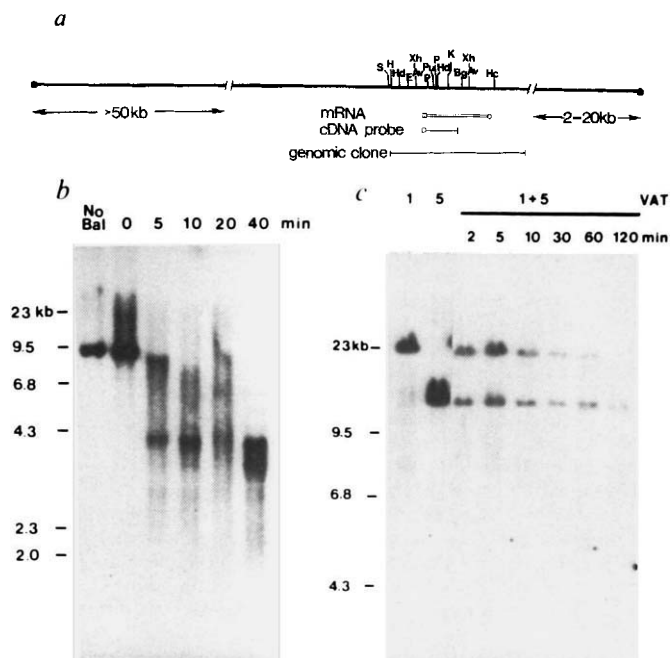
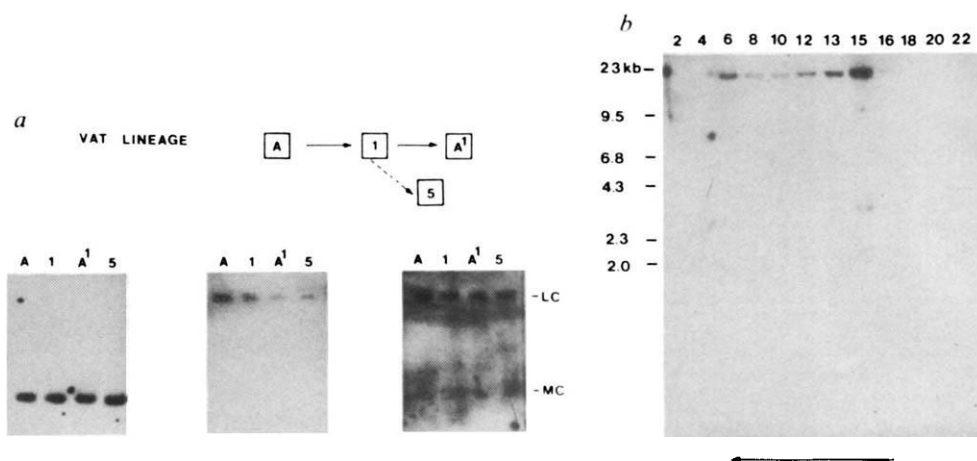


Fig. 1 Characteristics of the IsTat 1.1 VSG gene. **a**, Restriction map showing locations of the mature mRNA (open bar), cDNA probe and genomic clone (Av, *Ava*I; Bg, *Bgl*I; E, *Eco*RI; Hc, *Hinc*II; Hd, *Hind*III; K, *Kpn*I; P, *Pst*I; Pu, *Pvu*II; S, *Sal*I; Xh, *Xho*I; open box, SL; open circle, poly(A)). **b**, *Bal*31 nuclease sensitivity of the 1 VSG gene. VAT-11 (ref. 10) DNA was incubated with *Bal*31 according to Yao *et al.*²⁸ using 1 U of enzyme per 3 µg DNA; 5 µg DNA was taken per time point, deproteinized and then precipitated with ethanol. This DNA was digested with *Sal*I and transferred to nitrocellulose filters after electrophoretic separation in 0.8% agarose. The filters were probed with the nick-translated 1 VSG cDNA probe indicated under the restriction map **a**. λ DNA digested with *Hind*III was used as a molecular size standard. **c**, DNase I sensitivity of the 1 VSG gene. Equal numbers of VAT-1 and VAT-5 (ref. 10) cells were mixed, lysed and treated with DNase I for various times¹⁴. The DNA was purified, digested with *Sal*I, transferred to nitrocellulose after electrophoretic separation and hybridized with the cDNA probe. Lanes 1 and 5 contain DNA from these VATs which was not treated with DNase I. VAT nomenclature: the VAT (or expressed VSG) is indicated by the letter or number following the full stop (.).

expression and antigenic switching. Of several sequences associated with expressed VSG genes or their transcripts, only the 76-base pair (bp) repeat element was found in minichromosomal DNA. Variant antigenic types (VATs) which occurred immediately before and after VAT-1, expressed VSG genes located on larger chromosomes. Thus, different chromosomes can act as VSG gene expression sites.

We have shown previously¹⁰ that several VATs in the IsTaR 1 serodeme contain a single copy of the 1 VSG gene, whether or not it is expressed. We have also shown the presence of 1 VSG messenger RNA in VAT-1 (ref. 11). In all VATs examined, the restriction map is identical except for the length of the region 3' to the gene (Fig. 1a). Such length variation and the presence of an apparent cluster of restriction sites 3' to the gene are characteristic of telomeric VSG genes⁶. The telomeric location of the 1 gene was confirmed by its sensitivity to *Bal*31 exonuclease (Fig. 1b). The telomeric, but not the intrachromosomal, 11 VSG genes were sensitive to *Bal*31 (data not shown). That the single 1 VSG gene was activated without duplication was also confirmed by its preferential sensitivity to DNase I in VAT-1. In the experiment shown in Fig. 1c, VAT-1 and VAT-5 cells were mixed to allow a comparison of the DNase I sensitivity of the 1 VSG gene in cells expressing the 1 VSG with its sensitivity in cells expressing a different VSG. The 1 VSG gene in VAT-1 and VAT-5 cells can be distinguished by the length of their telomeres (also VAT-11, see Fig. 2b); the telomere is longer in

Fig. 2 Chromosomal locations of VSG genes. *a*, Freshly isolated trypanosomes were lysed in the wells of 0.5% agarose gels and transferred to nitrocellulose filters after electrophoretic separation according to Sloof *et al.*¹² and Allison *et al.*⁹. The filters were hybridized with nick-translated 1 VSG, A VSG or 5 VSG cDNA probes (left to right, respectively). The positions of the LC- and MC-DNA are indicated. Trypanosome clone 1 was isolated from the first relapse population from A, A¹ from the first relapse population from 1, and 5 from a relapse population occurring later in the same infection. A solid arrow indicates a single relapse, the broken arrow possible multiple relapses in the VAT lineage. Individual clones of each VAT are indicated by a subsequent lower case letter and the previous VAT (where known) by a superscript. All VATs used here belong to the IsTaR 1 serodeme and thus are more completely designated as 1.Ae, 1.1a, etc. For convenience the prefix (1.) and the lower case letter are omitted. *b*, Sucrose density gradient fractions of VAT-1 DNA prepared according to Allison *et al.*⁹ were digested with *Eco*RI and transferred to nitrocellulose filters after electrophoretic separation. The filters were hybridized with nick-translated 1 VSG cDNA probe. The lanes are marked with the fraction number and the arrow indicates the direction of sedimentation.



VAT-1 cells. The 1 VSG gene is more sensitive to DNase I in VAT-1 than in VAT-5 cells. Thus, these experiments show that the single 1 VSG gene is telomeric and was activated without duplication.

Although numerous restriction sites occur in and around the 1 VSG gene coding sequence, we could not find restriction sites 5' to the *Sal*I site (Fig. 1a), even using gels which resolve DNA fragments <50 kb in size. This distribution of restriction sites is similar to that which we⁹ and others⁸ have observed for VSG genes located on minichromosomes. Therefore, we investigated the location of the 1 VSG gene and other VSG genes in whole chromosome DNA separated into 'large' and minichromosomal-sized DNAs by gel electrophoresis¹² and by sucrose density gradient centrifugation⁸ (Fig. 2). The 1 VSG cDNA probe hybridized only to minichromosome-size DNA (MC-DNA) both in VATs expressing or not expressing the 1 gene (Fig. 2a), showing that this gene is on a minichromosome whether or not it is expressed.

We also examined the location of the A and 5 VSG genes. We have shown previously that there are two members of the A VSG gene family in the IsTaR 1 serodeme¹⁰, one telomeric and one intrachromosomal. The telomeric A gene is expressed without duplication in trypanosome clone A¹, derived directly from VAT-1. Clone A immediately preceded VAT-1 and contains an additional telomeric A VSG gene which has arisen by duplication of the other telomeric A VSG gene and is the transcribed expression-linked copy (ELC) expressed in trypanosome clone 5 (refs 10, 15). As the 5 VSG probe hybridizes to both MC- and LC-DNA in all VATs (Fig. 2a), the location of the transcriptionally active ELC itself is unknown.

When DNA was fractionated by sucrose gradient centrifugation, the 1 VSG cDNA probe hybridized primarily to the MC-DNA-containing fraction 15 (ref. 9 and Fig. 2b). Some hybridization to fractions 6–10 was also observed and probably results from MC-DNA trapped by the more abundant LC-DNA. As judged by ethidium bromide staining, most DNA was in fractions 2–4 where little or no hybridization with the 1 VSG probe occurred. Thus, our data provide further evidence that the 1 VSG gene is located on MC-DNA.

The activation of the 1 VSG gene on a minichromosome suggests that all elements required in *cis* for VSG gene expression and antigenic switching must be on at least one minichromo-

some. To investigate the identity of such sequences, we probed MC-DNA with the 76-bp repeat, the (TAA)*n* element and the spliced leader (SL) sequence. The 76-bp repeat and the (TAA)*n* element (which varies in size) have been found 5' to expressed¹⁶ and non-expressed^{15–18} VSG genes. The SL sequence occurs at the 5' end of VSG¹⁹ and other mRNAs^{20,21} and in tandem arrays and orphans in the genome²². Of these probes, only the 76-bp repeat hybridized to MC-DNA at high stringency (Fig. 3). At a lower stringency, the (TAA)*n* element also hybridized to MC-DNA. In preliminary studies^{7,21} and as will be reported in detail elsewhere, the SL probe did not hybridize to MC-DNA (R. G. Nelson *et al.*, in preparation) but all three probes hybridized to LC-DNA. Thus, the 76-bp repeat occurs on minichromosomes (although not necessarily the 1 VSG minichromosome), whereas an extensive (TAA)*n* element or SL sequence does not. The absence of these sequences from minichromosomes shows that they are not required in *cis* for VSG gene expression. These data leave open the possibility that the 76-bp repeat is required in *cis*, but, the presence of the 76-bp repeat 5' to unexpressed VSG genes^{15–18} indicates that it is not sufficient for VSG gene activation. Its location at or near the duplication boundary of VSG genes^{17,18} suggests that it has a role in duplication or transposition rather than gene activation *per se*.

We cloned from genomic DNA a segment containing ~1 kb of the 3'-flanking region of the 1 VSG gene (see Fig. 1a). The fragment 3' to the most 3' *Hinc*II site contains the repeat CCCTAA present on trypanosome telomeres (refs 23, 24 and R.A. Jr and K.S., unpublished results). This fragment hybridizes to numerous bands in genomic blots and to both MC- and LC-DNA, indicating that both types of chromosomes possess these telomeric sequences.

Our observations change the view of the significance of trypanosome minichromosomes. Williams *et al.*⁸ suggested that minichromosomal VSG genes are transitory intermediates in the production of an active VSG gene. The stability of the single minichromosomal 1 VSG gene through numerous cell divisions and clonings¹⁰ and through the entire life cycle of the trypanosome (R.A. Jr and K.S., unpublished results) indicates that minichromosomal VSG genes are neither transitory nor intermediate but contain all the sequences needed for faithful replication and segregation. Trypanosome minichromosomes are similar in size to the smallest chromosomes that can be propagated in yeast²⁵. These yeast chromosomes must contain telomeric and centromeric sequences for faithful segregation. Therefore, we postulate that minichromosomes contain centromeric sequences. Sloof *et al.*¹² proposed that minichromosomal VSG genes may evolve more rapidly than those on larger chromosomes, but there are insufficient data available to test

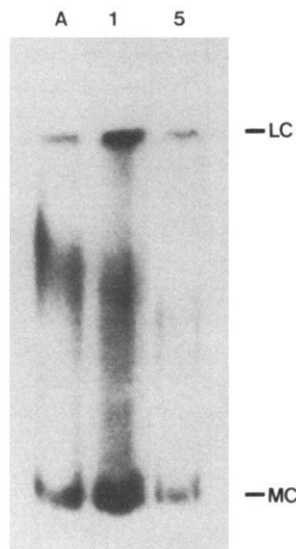


Fig. 3 Hybridization of 76-bp repeat probe with the MC- and LC-DNA. The filter of MC- and LC-DNA (prepared as described for Fig. 2a) was probed with a nick-translated 76-bp repeat probe. The probe was a *Sau3A/HinfI* fragment prepared from a clone of the 5' flank of the MITat 117 ELC and contains five 76-bp repeats¹⁵. The DNA used is indicated above each lane.

this hypothesis. All expressed VSG genes are telomeric and, as demonstrated here, are expressed from the telomeres of various chromosomes. As DNA organized into small chromosomes would have more telomeres than if organized into large chromosomes, minichromosomes may amplify the number of VSG gene expression sites. Additional evidence for multiple sites for expression of VSG genes has been presented by us^{9,14,20} and others⁷.

The expression of VSG genes on both minichromosomes and larger chromosomes shows that mechanisms other than a single expression site^{4,26} must control the exclusive expression of one of the numerous VSG genes. The 1 and A VSG genes in clones 1 and A¹ (A¹ was derived directly from 1, see Fig. 2 legend) were clearly not activated by reciprocal exchanges of telomeric segments containing these genes, as the active 1 VSG gene is on a minichromosome whereas the active A VSG gene is not. Although this does not eliminate telomeric exchange as one of the mechanisms for VSG gene activation²⁷, it shows that activation without duplication can occur by other mechanisms. The lack of differences in the restriction maps of the 1 VSG gene in its active and inactive states, although an imprecise criterion, provides no evidence for sequence addition, deletion or rearrangement correlated with VSG gene expression. Such differences, however, could exist far 5' to the gene or in the large region devoid of restriction sites, and therefore may not be detected. Thus, the identity and location of elements required in *cis* for the expression of VSG genes remains unresolved.

We thank John Scholler and Elke Gobrecht for technical assistance, Dr J. Allison for sucrose gradient fractions, Dr R. Nelson for communicating unpublished data and Dr J. Boothroyd for the MITat 117 ELC clone. This work was supported by NIH grants AI17309 and AI17375.

Received 4 September; accepted 20 November 1984.

1. Cross, G. A. M. *Parasitology* **71**, 393-417 (1975).
2. Hoeijmakers, J. H. J., Frasch, A. C. C., Bernards, A., Borst, P. & Cross, G. A. M. *Nature* **284**, 78-80 (1980).
3. Pays, E., Van Meirvenne, M., Le Ray, D. & Steinert, M., *Proc. natn. Acad. Sci. U.S.A.* **78**, 2673-2677 (1981).
4. Van der Ploeg, L., Bernards, A., Rijsewiyk, F. & Borst, P. *Nucleic Acids Res.* **10**, 593-609 (1982).
5. Young, J. R., Donelson, J. E., Majiwa, P. A. O., Shapiro, S. Z. & Williams, R. O. *Nucleic Acids Res.* **10**, 803-808 (1982).
6. DeLange, T. & Borst, P. *Nature* **299**, 451-453 (1982).
7. Van der Ploeg, L. H. T., Schwartz, D. C., Cantor, C. R. & Borst, P. *Cell* **37**, 77-84 (1984).
8. Williams, R. O., Young, J. R. & Majiwa, P. A. O. *Nature* **299**, 417-421 (1982).
9. Allison, J., Rothwell, V., Newport, G., Agabian, N. & Stuart, K. *Nucleic Acids Res.* **12**, 9051-9066 (1984).

10. Parsons, M. et al. *Molec. biochem. Parasit.* **9**, 255-269 (1983).
11. Milhausen, M. et al. *Molec. biochem. Parasit.* **9**, 241-254 (1983).
12. Sloof, P., Menke, H. H., Caspers, M. P. M. & Borst, P. *Nucleic Acids Res.* **11**, 3889-3901 (1983).
13. Myler, P., Nelson, R., Agabian, N. & Stuart, K. *Nature* **309**, 282-284 (1984).
14. Myler, P., Allison, J., Agabian, N. & Stuart, K. *Cell* (in the press).
15. Aline, R. Jr et al. *Nucleic Acids Res.* (submitted).
16. Campbell, D. A., VanBree, M. P. & Boothroyd, J. C. *Nucleic Acids Res.* **12**, 2759-2774 (1984).
17. Liu, A. Y. C., Van der Ploeg, L., Rijsewiyk, F. & Borst, P. *J. molec. Biol.* **167**, 57-75 (1983).
18. Van der Ploeg, L. et al. *Nucleic Acids Res.* **10**, 5905-5923 (1982).
19. Boothroyd, J. C. & Cross, G. A. M. *Gene* **20**, 281-289 (1982).
20. De Lange, T., Michels, P. A. M., Veerman, H. J. C., Cornelissen, A. W. C. A. & Borst, P. *Nucleic Acids Res.* **12**, 3777-3790 (1984).
21. Agabian, N. et al. *J. cell. Biochem.* (in the press).
22. Nelson, R. G. et al. *Cell* **34**, 901-909 (1983).
23. Blackburn, E. H. & Challoner, P. B. *Cell* **36**, 447-457 (1984).
24. Van der Ploeg, L. H. T., Liu, A. Y. C. & Borst, P. *Cell* **36**, 459-468 (1984).
25. Murray, A. W. & Szostak, J. W. *Nature* **303**, 189-193 (1984).
26. Borst, P. & Cross, G. A. M. *Cell* **29**, 291-303 (1982).
27. Borst, P. et al. in *Gene Expression* Vol. 8 (eds Hamer, D. H. & Rosenberg, M.) 413-435 (Liss, New York, 1983).
28. Yao, M. C. *Cell* **24**, 765-774 (1981).

Close link between reduction of *c-myc* expression by interferon and G₀/G₁ arrest

Miriam Einat, Dalia Resnitzky & Adi Kimchi

Department of Virology, The Weizmann Institute of Science, Rehovot, Israel

It has recently been reported that *c-myc* is an inducible gene, regulated directly by growth signals which promote proliferation and expressed in a cell-cycle dependent manner¹. Because various leukaemic cell lines express high levels of *c-myc* messenger RNA², it was of interest to discover whether the gene could be down-regulated in these cells by a growth inhibitor such as interferon (IFN)³. We show here that in Daudi Burkitt's lymphoma cells, IFN- α produces a five- to sevenfold reduction in *c-myc* mRNA through a decreased rate of *c-myc* gene transcription. By isolating a growth-resistant Daudi cell variant that had escaped from this down-regulation, we provide the first clear link between reduction of *c-myc* mRNA and the IFN-mediated G₀/G₁ arrest characteristic of Daudi cells. Furthermore, by screening other cell lines, we demonstrate the heterogeneity of human leukaemic cells with respect to these criteria. Thus, IFN- α fails to reduce the *c-myc* mRNA and to change the cell-cycle distribution in HL-60 and U937 cells, although normal induction of other IFN-regulated activities takes place. The latter group of cells shows a decline in *c-myc* gene expression when they become arrested in the G₀/G₁ phase as part of their terminal differentiation.

Three leukaemia cell lines, HL-60 (ref. 4), U937 (ref. 5) and Friend erythroleukaemia⁶, representing different cell lineages of the haematopoietic system, were induced to differentiate by chemical agents such as dimethyl sulphoxide (DMSO) or phorbol myristate acetate (PMA) (Fig. 1 legend). Cytofluorimetric analysis was performed on these cultures before and after induction of differentiation and the proportion of cells in the different phases of the cell cycle was assessed (Fig. 1a). In parallel experiments, equal amounts of poly(A)⁺ RNA were analysed for *c-myc* mRNA levels by Northern transfer and hybridization to nick-translated *c-myc* DNA sequences (Fig. 1b). A single band of ~2.4 kilobases (kb) was detected in the different leukaemic cells by the *c-myc* probe.

Figure 1 shows that in HL-60, U937 and Friend erythroleukaemia cells, terminal differentiation triggered by PMA or DMSO was accompanied by two common events, an increase in the proportion of cells in the G₀/G₁ phase at the expense of the S, G₂ and M phases (defined as a G₀/G₁ arrest), and a sharp decline in the amount of *c-myc* mRNA. A fourth cell line, Daudi Burkitt's lymphoma cells⁷, failed to respond to PMA according to these parameters.

To analyse the possible relationships between these two events, we exposed the leukaemic cell lines to a growth inhibitor, 100 U ml⁻¹ of IFN- α purified to homogeneity by monoclonal antibody columns as described elsewhere⁸. Figure 1b shows that

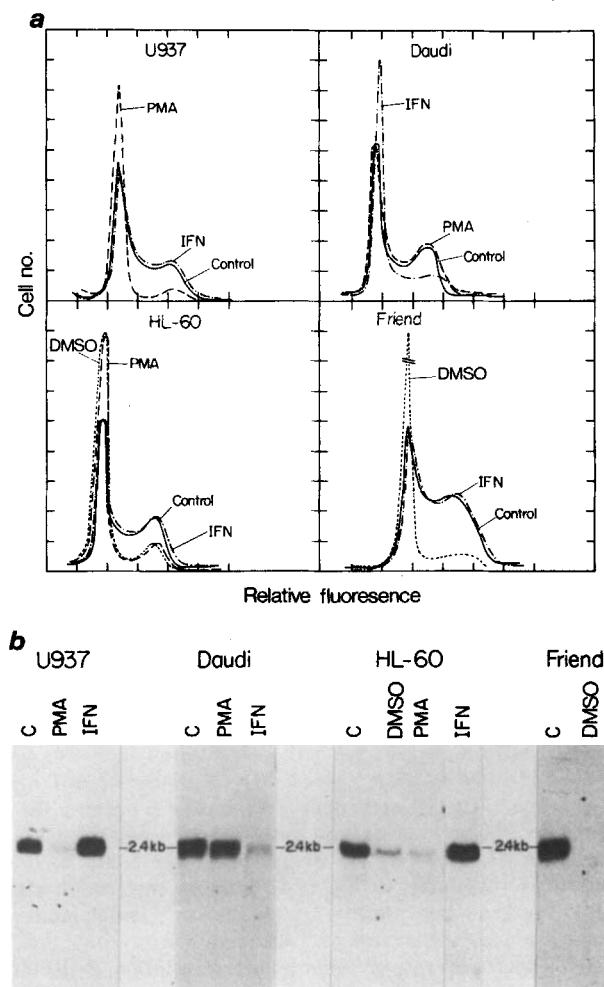


Fig. 1 Growth arrests of leukaemic cell lines mediated by IFN- α or by chemical inducers of cell differentiation. *a*, Cytofluorimetric analysis of control non-treated cells (—) and of cells treated with either PMA (—), DMSO (----) or IFN (— · —). *b*, Northern blot hybridization analysis of *c-myc* mRNA from control cells (C) and from cells treated as in *a*.

Methods. The human histiocytic lymphoma U937, promyelocytic leukaemia HL-60 and Daudi Burkitt's lymphoma cell lines were grown in RPMI medium (Gibco) supplemented with 10% fetal calf serum, and treated for 48 h with 100 U ml⁻¹ IFN- α at a cell density of 2×10^5 cells ml⁻¹. Induction of differentiation of U937 and HL-60 towards macrophages was performed by exposing 2×10^5 cells ml⁻¹ to 50 ng ml⁻¹ PMA¹⁰ (Sigma), and of HL-60 cells to granulocytes using 1.25% DMSO⁴; cells were analysed 48 h post-induction. Friend erythroleukaemia cells were grown and induced by 1.5% DMSO as we have described previously in detail¹³, and analysed on day 7 post-induction, or 48 h post-treatment with 100 U ml⁻¹ murine IFN¹³. Cytofluorimetric analysis following propidium iodide staining was performed as described previously¹⁴. Total RNA was prepared by the LiCl method from $1-5 \times 10^8$ cells as detailed elsewhere⁹, and poly(A)⁺ RNA was selected by chromatography over oligo(dT)-cellulose. Poly(A)⁺ RNA from U937, Daudi and Friend cells (10 μ g) from HL-60 cells (3 μ g) was gel electrophoresed and transferred to nitrocellulose sheets as described elsewhere¹⁰. Hybridization to 10⁷ c.p.m. of the *c-myc* probe radiolabelled by nick translation to a specific activity of 10⁸ c.p.m. μ g⁻¹ was performed as before¹⁰. The *c-myc* probe was a 1.4-kb *Clal*/*EcoRI* DNA fragment containing the third exon of the human *c-myc* gene (provided by R. Gallo)¹⁵.

in the Daudi cell system, IFN significantly reduced the amounts of *c-myc* mRNA; in contrast, *c-myc* mRNA levels were stimulated by a similar exposure of U937 and HL-60 cells to IFN. Interestingly, IFN induced a typical arrest in G₀/G₁ only in Daudi cells (Fig. 1a) whereas U937, HL-60 or Friend erythroleukaemia cells showed no change in the cell-cycle distribution, although cell growth was reduced by IFN (Table 1).

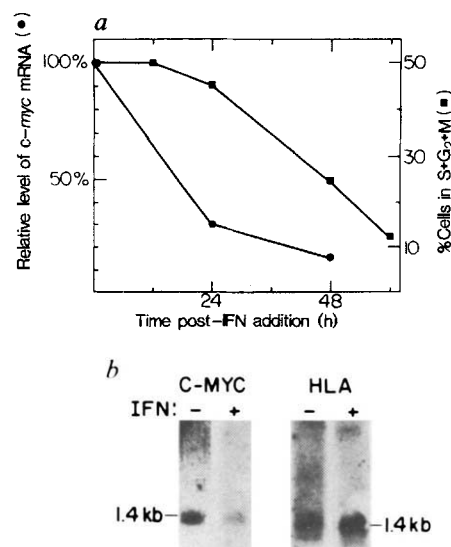


Fig. 2 Down-regulation of the *c-myc* gene by IFN- α in Daudi cells. *a*, Graphic representation of relative levels of *c-myc* mRNA (●), and of the percentage of cells in S+G₂+M replicative phases (■) at various time points after exposure to IFN- α . *b*, Autoradiograms of hybridization of *in vitro* transcribed RNA to DNA probes. Nuclei were isolated from control non-treated Daudi cells (—) or from cells pretreated for 20 h with 100 U ml⁻¹ IFN- α (+).

Methods. *a*, Daudi cells were grown and exposed to 100 U ml⁻¹ IFN- α and at various times later they were prepared for cytofluorimetry and the proportion of cells in S, and in G₂+M was determined from the histograms as described previously¹⁶. RNA was extracted from the cells at the indicated times and 5 μ g poly(A)⁺ RNA were analysed on Northern blots. The relative levels of *c-myc* mRNA were determined from densitometer scanning of the autoradiograms after different exposure times. *b*, *In vitro* synthesis of *c-myc* mRNA in nuclei. Nuclei were isolated from 10⁸ Daudi cells and disrupted by 20–30 strokes of a Dounce homogenizer in 40 ml of 10 mM HEPES-KOH pH 7.6, 5 mM MgCl₂, 25 mM KCl, 0.1 mM EDTA, 1 mM dithiothreitol (DTT) (hypotonic buffer). The 4,000g pellet was suspended in 5 ml of hypotonic buffer, recentrifuged for 8 min through 10 ml of 25% glycerol in the same buffer, resuspended at 10⁸ nuclei ml⁻¹ in the above glycerol buffer and stored in liquid air. Nuclei (2×10^7) were incubated for 30 min at 25 °C in 0.3 ml of 25 mM HEPES-KOH pH 7.5, 150 mM KCl, 6 mM MgCl₂, 1 mM MnCl₂, 0.5 mM K₂HPO₄, 0.03 mM EDTA, 5 mM DTT, 5 mM NaF, 5% glycerol, 25 μ M UTP, 250 μ Ci [α -³²P]UTP (400 Ci mmol⁻¹, Amersham) and 0.4 mM of GTP, CTP and ATP. The nuclei were then treated with 30 μ g ml⁻¹ DNase I for 5 min (Worthington, RNase free). RNA was phenol and chloroform extracted, treated with DNase I (10 μ g ml⁻¹), re-extracted by phenol/chloroform and filtered through Sephadex G-50. Recombinant plasmids containing the 1.4-kb genomic fragment of *c-myc* DNA (see Fig. 1) and the 1.4-kb HLA-B cDNA¹⁷ were cleaved by restriction enzymes to generate the specific inserts, electrophoresed on 1% agarose gels (2 μ g per slot) and blotted onto nitrocellulose¹⁸. Blots (1 cm strip) were incubated for 4 h at 42 °C in 1 ml of 50% formamide, 0.1 M HEPES-KOH pH 7.5, 0.75 M NaCl, 1 mM EDTA, 0.2% polyvinylpyrrolidone, 0.2% Ficoll and 0.5% SDS before addition of 5×10^6 c.p.m. of ³²P-labelled nuclear RNA for 48 h. Blots were washed twice in 2 $\times$ SSC, treated for 30 min at room temperature with 20 μ g ml⁻¹ RNase A, washed for 1 h at 42 °C in 2 $\times$ SSC, 0.5% SDS, three times for 30 min at 50 °C in 0.5 $\times$ SSC, 0.5% SDS, and exposed for 2 days.

Densitometer tracing of the autoradiograms obtained from several experiments indicated a five- to sevenfold reduction in *c-myc* mRNA level in IFN-treated Daudi cells, compared with a two- to threefold increase in IFN-treated U937 and HL-60 cells. We therefore conclude that the various leukaemic cell lines differ in their response to IFN with respect to the effect on the *c-myc* mRNA levels and the development of a specific G₀/G₁ arrest. No significant difference between these cell lines was detected, however, in the extent of induction of (2'-5') oligoadenylate (oligo(A)) synthetase by IFN, although the basal

and maximal levels of synthetase activity were higher in Daudi cells than in the other cell lines (Table 1).

Further analysis of the inhibitory effect of IFN on *c-myc* gene expression showed that the IFN- α -induced decline in *c-myc* mRNA in Daudi cells preceded the IFN-mediated reduction of cells in S+G₂+M phases (Fig. 2a). To determine whether the down-regulatory effect of IFN operates at transcriptional or post-transcriptional levels, we measured the *c-myc* gene transcription in isolated nuclei. We found that 5×10^6 c.p.m. of *in vitro* synthesized ³²P-labelled nuclear RNA were hybridized to the 1.4-kilobase (kb) DNA fragment of the *c-myc* gene blotted onto nitrocellulose (Fig. 2b). As an internal control, the same preparations of *in vitro* transcribed RNA were hybridized to the HLA-B cDNA probe. Whereas nuclei isolated from cells pre-treated for 20 h with IFN transcribed *c-myc* mRNA much less efficiently than nuclei from control untreated cells (Fig. 2b), the transcriptional rate of HLA mRNA in the same preparations was significantly increased. IFN therefore simultaneously reduced and activated the transcription rate of two different sets of genes; this could lead to the observed changes in the steady-state levels of mRNA, that is, the decline in *c-myc* mRNA and the previously described increase in HLA class I mRNA^{9,10}. In another experiment, a reduction in *c-myc* transcription was detected within 10 h of IFN treatment (not shown), suggesting that the effect of IFN on *c-myc* expression is not secondary to changes in cell-cycle distribution (compare with Fig. 2a).

We isolated a variant of Daudi cells resistant to the inhibitory effect of 100 U ml⁻¹ IFN- α (IFN^r cells) (Fig. 3 legend) but whose growth was markedly reduced at higher IFN- α concentrations (1,000 U ml⁻¹) (Table 1). Cytofluorimetric analysis indicated that these cells no longer showed IFN-mediated G₀/G₁ arrest, that is, any reduction in the proportion of cells in S+G₂+M phases, even with high concentrations of IFN (1,000 U ml⁻¹) (Fig. 3, Table 1). Figure 3 also shows that, in contrast to the reduction seen in wild-type cells, *c-myc* mRNA levels doubled in IFN^r cells following exposure to IFN- α . Hybridization of the same blots with the (2'-5')oligo(A) synthetase E cDNA probe recently isolated by Merlin *et al.*¹¹ revealed a strong induction of the synthetase 1.8-kb mRNA by IFN in the two cell lines (Fig. 3). As expected, synthetase activity also increased in response to IFN in these cells (Table 1). The IFN^r cells are therefore not defective in cell-surface receptors for IFN- α ; they respond to IFN with respect to the (2'-5')oligo(A) synthetase induction, but have specifically lost the

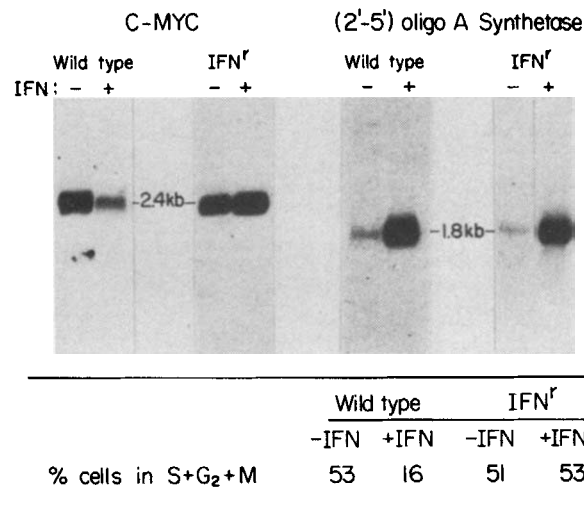


Fig. 3 IFN-resistant (IFN^r) Daudi cells have lost both the G₀/G₁ arrest and the down-regulation of the *c-myc* gene by IFN but not the IFN-mediated increase of (2'-5')oligo(A) synthetase mRNA. Northern blot hybridization analysis of *c-myc* mRNA and of (2'-5')oligo(A) synthetase mRNA in non-treated (-) and IFN-treated (+) Daudi wild-type and IFN^r cells. Also shown is the fraction of cells in S+G₂+M phases determined from cytofluorimetry histograms as in Fig. 2a.

Methods. Daudi wild-type and IFN^r cells were grown and treated with 100 U ml⁻¹ IFN- α for 48 h as in Fig. 1. Poly(A)⁺ RNA (8 μ g) was analysed for *c-myc*-specific sequences; blots were then stripped of the *c-myc* probe by boiling for 2 min in H₂O, and hybridized with the cloned (2'-5')oligo(A) synthetase E cDNA probe, recently isolated¹¹. Selection of the IFN^r cells: Exponentially growing Daudi cells were treated with 100 U ml⁻¹ of IFN- α and the growth-arrested cultures were kept for several weeks until resistant cells capable of growing in the presence of IFN appeared. IFN^r cells were continuously maintained with 100 U ml⁻¹ IFN- α ; before the experiment they were grown for 10-20 days without IFN.

Table 1 Induction of (2'-5')oligo(A) synthetase and growth inhibition by IFN- α in leukaemic cell lines

Cells	IFN (U ml ⁻¹)	(2'-5')oligo(A) synthetase (c.p.m. per 10 μ g protein)	Growth % inhibition	Reduction in S+G ₂ +M
U937	0	1,600		
	100	25,800	34	-
	1,000	22,000	66	-
HL-60	0	1,400		
	100	19,900	13	-
Daudi (wild type)	0	6,500		
	10	27,100		+
	100	82,300	70	+
	1,000	70,400	70	+
Daudi (IFN ^r)	0	5,200		
	10	28,100		-
	100	40,100	10	-
	1,000	44,400	56	-

Exponentially growing cells (2×10^5 cells ml⁻¹) were treated with IFN- α and 24 h later cell extracts were prepared and samples tested for (2'-5')oligo(A) synthetase activity as described previously¹⁰. Cells were counted at 72 h and % reduction of cell number by IFN was calculated. The ability of IFN to reduce the percentage of cells in S+G₂+M phases (+) was determined from cytofluorimetric analysis. IFN^r cells were grown for 20 days without IFN before the experiments.

IFN-mediated down-regulation of the *c-myc* gene and the development of a G₀/G₁ arrest.

Jonak *et al.*¹² reported recently that IFN- β reduces *c-myc* mRNA levels in Daudi cells and that this reduction is selective because neither the level of actin mRNA nor the total cellular protein synthesis is affected. We have now shown that this alteration results from a reduction in the transcription rate of the *c-myc* gene, and by comparing the wild-type cells with an IFN^r Daudi cell variant, we have further established a tight connection between the decline of *c-myc* mRNA induced by IFN and the subsequent IFN-mediated G₀/G₁ arrest. Based on this correlation, we propose that the quantitation of *c-myc* mRNA should be used to predict the sensitivity of fresh peripheral leukaemic cells to IFN.

The basis of the difference between Daudi cells and other leukaemic cell lines such as HL-60 and U937, in which IFN failed to reduce *c-myc* mRNA, is unclear. HL-60 and U937 cells display normal induction of other IFN-regulated activities and it is possible that these cells have specifically escaped from the inhibitory effect of IFN on the *c-myc* gene during the development of the leukaemia. Within the group of Burkitt's lymphoma cells, some cell lines, such as Namalva and Raji, are also resistant to IFN by these criteria, whereas the MOLT-4 cells respond like Daudi cells (not shown). It is clear, however, that the availability of these resistant leukaemic cell lines, as well as the IFN^r Daudi cell variant, provides new approaches to investigating the molecular mechanism which might lead to this type of *c-myc* deregulation.

We thank Dr R. Gallo for the gift of the *c-myc* probe, Dr M. Rubinstein for purified IFN- α and Drs D. G. Fischer and M. Revel for helpful discussions and for providing the (2'-5')oligo(A) synthetase E cDNA probe. This work was supported by the Leukemia Research Foundation, Chicago. A.K. is an incumbent of the Louis and Ida Rich Career Development Chair.

Received 7 September; accepted 7 November 1984.

- Kelly, K., Cochran, B. H., Stiles, C. D. & Leder, P. *Cell* **35**, 603-610 (1983).
- Westin, E. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2490-2494 (1982).
- Taylor-Papadimitriou, J. in *Interferon* Vol. 2 (ed. Gresser, I.) 13-46 (Academic, New York, 1980).
- Collins, S. J., Ruscetti, F. W., Gallagher, R. E. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2458-2462 (1978).
- Sundström, C. & Nilsson, K. *Int. J. Cancer* **7**, 565-577 (1976).
- Marks, P. A. & Rifkind, R. A. *Rev. Biochem.* **47**, 419-448 (1978).
- Klein, E. *et al. Cancer Res.* **28**, 1300-1310 (1967).
- Novick, D., Eshhar, Z. & Rubinstein, M. *J. Immun.* **129**, 2244-2247 (1982).
- Fellous, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3081-3086 (1982).
- Yarden, A., Shure-Gottlieb, H., Chebath, J., Revel, M. & Kimchi, A. *EMBO J.* **3**, 969-973 (1984).
- Merlin, G., Chebath, J., Benech, P., Metz, R. & Revel, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4904-4908 (1983).
- Jonak, G. J. & Knight, E. Jr *Proc. natn. Acad. Sci. U.S.A.* **81**, 1747-1750 (1984).
- Friedman-Einat, M., Revel, M. & Kimchi, A. *Molec. cell. Biol.* **2**, 1472-1480 (1982).
- Bendori, R., Resnitsky, D. & Kimchi, A. *FEBS Lett.* **162**, 384-389 (1983).
- Dalla-Favera, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6497-6501 (1982).
- Cullen, M. H. & Capellaro, D. F. in *Pulse-Cytophotometry: Proc. 3rd int. Symp.* (ed. Lutz, D.) 643-657 (European Press, Ghent, 1978).
- Sood, A. K., Pereira, D. & Weissman, S. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 616-620 (1981).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).

Specific expression of an elastase-human growth hormone fusion gene in pancreatic acinar cells of transgenic mice

David M. Ornitz*, Richard D. Palmiter*, Robert E. Hammer†, Ralph L. Brinster†, Galvin H. Swift‡ & Raymond J. MacDonald‡

* Howard Hughes Medical Institute Laboratory, Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

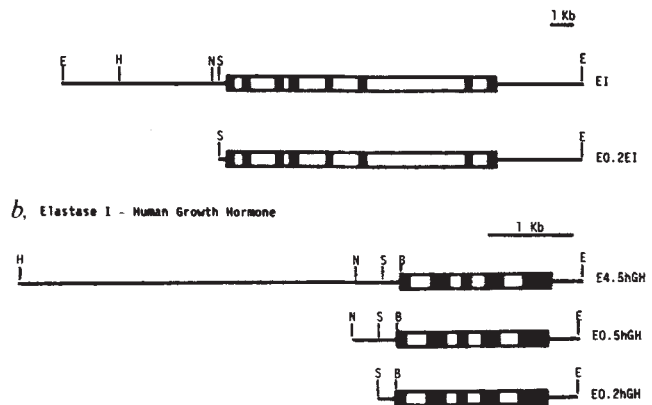
† Laboratory of Reproductive Physiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

‡ The Division of Molecular Biology, Department of Biochemistry, The University of Texas Health Science Center at Dallas, Dallas, Texas 75235, USA

Transfection of genes into tissue culture cell lines has demonstrated that relatively short DNA sequences can allow expression of immunoglobulin¹⁻³, insulin and chymotrypsin⁴ genes in their appropriate cell types. A definitive test of cell-specific gene expression, however, requires testing genes in every possible cell type, an experiment performed easily by introducing the gene in question into the germ line of an animal. Transfer of intact genes into mice has demonstrated that a mouse immunoglobulin κ gene is expressed specifically in B lymphocytes^{5,6}, a rat elastase I gene is expressed specifically in pancreas⁷ and a chicken transferrin gene is expressed preferentially in liver⁸. Mouse metallothionein-growth hormone fusion genes introduced into mice are preferentially expressed in the liver, consistent with the expression of endogenous metallothionein genes⁹, but initial experiments with β -globin genes have not revealed proper regulation¹⁰⁻¹². To identify the DNA elements required for pancreas-specific expression of the rat elastase I gene, we joined the 5'-flanking region of this gene to the human growth hormone (hGH) structural gene and introduced the fusion gene into mice. Here we demonstrate that a fusion gene containing only 213 base pairs (bp) of elastase I gene sequence directs expression of hGH in pancreatic acinar cells.

Figure 1a, b shows: (1) the elastase I gene which functions in transgenic mice⁷; (2) the truncated version tested here; (3) three elastase-hGH fusion genes containing 4.5, 0.5 and 0.2 kilobases (kb) of elastase 5'-flanking sequence. The elastase-hGH fusion gene was constructed by placing a *Bam*HI linker at the +8 position of the elastase I gene and ligating it to the *Bam*HI site at the +2 position of the hGH gene¹³. The resulting messenger RNA derived from this fusion gene should contain the first 8 nucleotides of the elastase I mRNA, followed by 2 nucleotides contributed by the synthetic linker, then hGH sequence starting at +2 (Fig. 1c). The hGH structural gene has several advantages as a marker gene in transgenic mice. The

a, Elastase I



c, Elastase I - Human Growth Hormone Sequence

```

-210 CTGAGTCGAC TTGGGTTAAG TGAGTCCGG CCTTGTCTG TCTTTGAATA TCAGATAAAT GAGTGTACTT
-140 AAAATTGTGT CATTGTGACT TCGATGTCAC CTGTGCTTTT CCGTCCCTTC TACCCACAGC CCGGCCAGCC
-70 TGCCGAGGGA AGCTCAGCAG AGCTGTGAT AAGAGCCGTA TAAGAGGGT TCCGTCATG SCAGGGGCC
+1 AGTGTCTCC GGATCCAG GCCCACTC CGAAGCACT CAGGCTCTG TGACAGCTC ...NGH...

```

Fig. 1 Rat elastase I and elastase-human growth hormone gene constructions. a, The rat elastase I gene (EI) and the 5' deletion at -205 (E0.2EI) which was introduced into mice as a linear *Sall*/*Nru*I fragment containing 3.4 kb of pBR322 vector sequence (see ref. 7). b, The three elastase-hGH fusion genes that were introduced into mice. The fusion genes were constructed by subcloning a 403-nucleotide *Sall*/*Stu*I fragment spanning -205 to +198 of rat elastase into pUC13, then converting an *Acc*I site at +8 to a *Bam*HI site with a synthetic linker. The hGH structural gene, contained on a *Bam*HI/*Eco*RI fragment¹³, was then inserted to make pE0.2hGH. pE4.5hGH was constructed by inserting the 4.3-kb 5' elastase *Hind*III/*Sall* fragment into pE0.2hGH. E0.2hGH and E4.5hGH were injected as *Hind*III linear fragments that included 2.6 kb of pUC13 vector sequence, whereas E0.5hGH was injected as an *Nde*I fragment, excised from pE4.5hGH, containing 0.2 kb of vector sequence. Exons are shown as solid boxes, introns as open boxes. E, *Eco*RI; H, *Hind*III; N, *Nde*I; S, *Sall*. c, Sequence of the elastase DNA present in E0.2hGH. The putative tissue-specific enhancer region and TATAAA box are over- and underlined. The sequence to the right of +10 (underlined) corresponds to the hGH structural gene.

presence of hGH DNA and mRNA can be assayed easily by hybridization with cDNA or synthetic oligomers. Human GH mRNA can accumulate to a high steady-state concentration in a variety of mouse tissues and if it is translated appropriately and secreted into the blood, increased growth should be observed⁹.

Each of the plasmids to be tested was linearized and micro-injected into the male pronucleus of fertilized mouse eggs (F₂ hybrids of C57 × SJL crosses), which were then implanted into pseudopregnant, surrogate mothers. Out of 131 pups derived from microinjected eggs, dot hybridization of tail DNA identified 4 mice with the elastase I gene and 18 mice with elastase-hGH fusion genes (Table 1). The mice were monitored for growth for at least 8 weeks and then killed for tissue analysis. Rat elastase and hGH mRNA were assayed by a convenient solution hybridization method using either a labelled 21-base oligomer or a complementary DNA probe (see Table 1 legend). All four mice containing the truncated gene (E0.2EI) expressed rat elastase mRNA in the pancreas (Table 1a). Of the 18 transgenic mice with the fusion genes, 5 had no detectable hGH mRNA, 1 had a very low concentration of hGH mRNA in the pancreas and the remaining 12 had 1,100-39,000 molecules of hGH mRNA per pancreatic cell (Table 1b). The data obtained with the truncated elastase I gene and the E0.2hGH fusion gene indicate that 213 bp of elastase promoter region (Fig. 1c) are sufficient to direct pancreas-specific expression.

All of these transgenic mice had <10 molecules of hGH mRNA per cell in the non-pancreatic tissues examined, which included intestine, kidney, liver, parotid, spleen, submandibular

Table 1 Expression of rat elastase and hGH in tissues of transgenic mice

Plasmid	Mouse	Genes per cell	mRNA	
			Pancreas (molecules per cell)	Other tissues
a, Elastase				
E0.2EI	26-1	13.0	1,900	<10
	66-2	9.5	3,600	<10
	55-6	47.0	3,900	<10
	62-1	51.0	4,100	<10
b, hGH				
E4.5hGH	37-9	6.5	0	<10
	31-7	2.2	35	<10
	34-4	1.5	1,170	<10
	34-9	3.2	2,560	<10
	34-2	1.4	4,000	<10
	35-1	1.2	8,490	<10
	33-6	6.2	15,100	<10
	34-10	6.2	28,100	<10
E0.5hGH	43-4	1.0	0	<10
	43-10	6.0	0	<10
	42-6	1.2	1,460	<10
	44-6	199.0	9,760	<10
	44-4	136.0	11,900	<10
	43-5	2.3	15,400	<10
	40-2	4.2	39,400	<10
E0.2hGH	47-4	3.3	0	<10
	48-4	39.4	0	<10
	49-3	2.1	3,340	<10

To calculate mRNA molecules per cell, we used the following values: 6.4 pg DNA per cell, hGH mRNA averages 950 bases, elastase mRNA averages 1,200 bases and the RNA/DNA ratios are: pancreas, 8.4; intestine, 1.5; kidney, 0.96; liver, 5.3; parotid, 2.2; spleen, 0.38; submandibular gland, 2.2; testes, 1.75; uterus, 2.0. The limit of sensitivity of the assay is ~10 molecules per cell. The number of genes per cell was determined by quantitative DNA dot hybridization using an hGH probe. Hybridization to the endogenous mouse metallothionein-I gene was used as an internal control.

Methods. mRNA levels were quantitated using the solution hybridization protocol described by Durnam and Palmiter¹⁶. For elastase, RNA was prepared by the guanidine thiocyanate procedure¹⁷ and the probe was a single-stranded cDNA specific for rat elastase I, prepared as described by Ashley and MacDonald¹⁸. For hGH mRNA, total nucleic acids were prepared by the SDS-proteinase K method¹⁶ and the probe was a synthetic 21-nucleotide oligomer (5' ATAGACGTTGCTGTCAGAGGC 3') complementary to a sequence present in the last exon. The oligomer was end-labelled with [γ -³²P]ATP to a specific activity of ~3,000 Ci mmol⁻¹ by incubating 20 pmol of labelled ATP and 10 pmol of oligomer for 30 min at 37 °C in a 10- μ l reaction with T4 polynucleotide kinase. The labelled oligomer was isolated on a 12% acrylamide gel and eluted in 0.2 $\times$ SET (1 $\times$ SET = 1% SDS, 5 mM EDTA, 10 mM Tris-HCl, pH 7.5). The hybridization reactions were set up exactly as described elsewhere¹⁹, except that 5,000 c.p.m. of labelled oligomer were used per reaction, formamide was replaced with 0.2 $\times$ SET, the dilution buffer for all of the samples and controls included 100 μ g ml⁻¹ of herring sperm DNA in 0.2 $\times$ SET, hybridization temperature was lowered to 45 °C, the S₁ nuclease treatment was at 37 °C and the trichloroacetic acid precipitates were allowed to sit on ice for 1 h before they were collected on GF/C filters. A standard curve was established using M13 single-stranded DNA containing the hGH structural gene¹³. The use of labelled oligomers obviates the cumbersome preparation of cDNA. Using the modifications described above, it is possible to detect >10 mRNA molecules per cell with a very low background. Furthermore, oligomers can be designed to discriminate between closely related sequences.

and testis or uterus. In most samples, no hGH mRNA was detected. The low concentrations detected occasionally are probably insignificant because none of the mice showed increased growth, indicating the lack of secretion of significant amounts of a functional hGH protein into the bloodstream. Many of the transgenic mice had at least 10³–10⁴-fold more hGH mRNA in the pancreas than in any other tissue analysed. We have noted that both gene constructs with 205 bp of elastase 5'-flanking sequence gave lower levels of expression than the corresponding

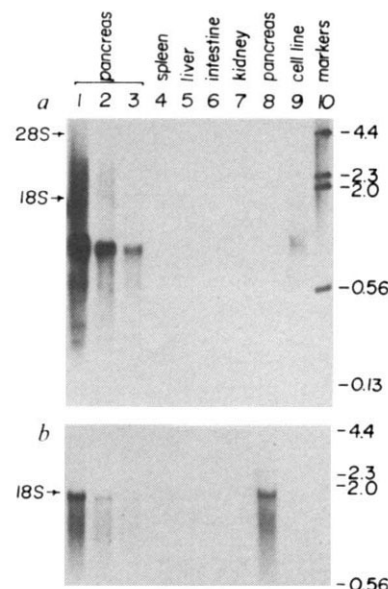


Fig. 2 Size of hGH mRNA and tissue specificity as determined by Northern blot hybridization. *a*, Blot analysed with a ³²P-labelled nick-translated hGH probe; *b*, the same blot re-analysed with a kinase-treated ³²P-oligomer specific for mouse pancreatic amylase. Lanes 1–3 contain 10, 3 and 1 μ g, respectively, of pancreatic RNA from an offspring of mouse 40–2. Lanes 4–7 contain 10 μ g of RNA from spleen, liver, intestine and kidney, respectively. Lane 8 contains 10 μ g of RNA from a control mouse pancreas and lane 9 contains 6 μ g total nucleic acid from a hepatocellular carcinoma cell line expressing metallothionein-hGH. Lane 10 contains end-labelled λ DNA digested with *Hind*III; arrows indicate 18S and 28S RNA bands. Staining the nitrocellulose with methylene blue confirmed that equal amounts of RNA were loaded in lanes 1 and 4–8.

Methods. Total RNA was prepared by homogenizing tissue in 5 M guanidine thiocyanate, 1% sarkosyl, 20 mM EDTA, 1% 2-mercaptoethanol, 50 mM Tris pH 7.5, followed by centrifugation through 5.7 M CsCl^{17,19}. RNA (10 μ g) and DNA standards were denatured in 50% formamide, 2.2 M formaldehyde at 68 °C for 5 min and electrophoresed at 200 V through a 1.5% agarose gel containing formaldehyde²⁰. Gels were blotted onto nitrocellulose and hybridized by standard procedures.

genes with more 5'-flanking sequence. Because there is an imperfect relationship between gene copy number and expression, we cannot be certain whether sequences upstream of ~205 contribute to the level of gene expression in the pancreas. The proportion (5/18) of transgenic mice which failed to express the fusion gene is comparable with that observed with metallothionein-hGH⁹ and presumably represents integration into an inactive chromatin domain. This interpretation suggests that chromatin domains are dominant over both elastase and metallothionein promoters, even though the former promoter is tissue-specific and capable of producing much higher levels of hGH mRNA than the latter⁹.

When offspring of transgenic mice were analysed for elastase-hGH expression, we found that offspring of mouse 43–5, carrying twice the parental gene copy number, expressed higher concentrations of hGH mRNA (53,400 molecules per cell) in the pancreas, which indicates that the parent is probably mosaic and that pancreas-specific expression of hGH can be transmitted to the next generation. Similarly, pancreas-specific expression of rat elastase has been transmitted to offspring in five separate lines (G.H.S. and R.J. MacD., unpublished observations). Thus, expression of these genes, like the metallothionein-hGH fusion genes¹⁴, is usually transmitted faithfully to succeeding generations.

Northern blot analysis was used to investigate the processing of elastase-hGH mRNA in tissues of transgenic mice. Figure 2*a* confirms that hGH mRNA is present only in pancreatic tissue. The hGH mRNA is ~950 nucleotides, the same size as

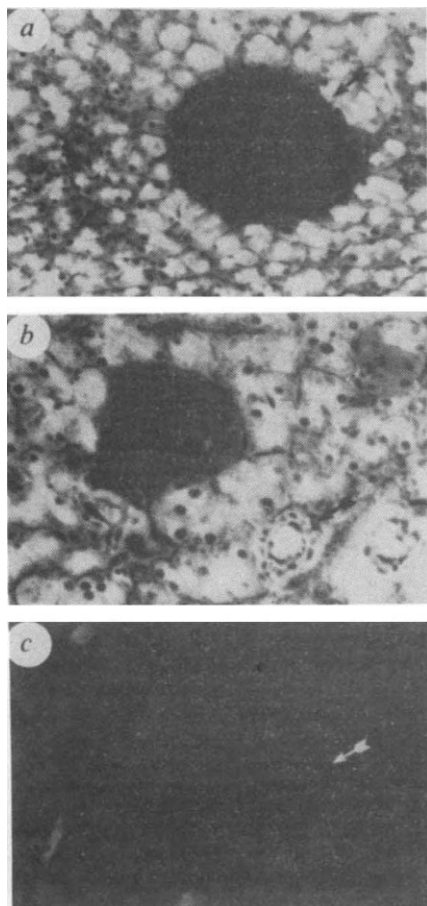


Fig. 3 Immunofluorescent localization of human growth hormone. *a*, A section of pancreas from mouse 34-10 treated with rabbit anti-hGH and then with fluorescein isothiocyanate (FITC)-labelled, goat anti-rabbit immunoglobulin. The exocrine acinar cells fluoresce brightly while the endocrine islet cells (arrow) are dark. *b*, A higher magnification view of pancreas from mouse 34-10 demonstrating hGH located in the acinar cells and in a pancreatic duct (arrow). *c*, Pancreas of the same mouse treated with normal rabbit serum before FITC-labelled, goat anti-rabbit immunoglobulin. The arrow indicates islet cells.

Methods. Immediately after the mouse had been killed, the pancreas was dissected and immersed in Carnoy's fixative for at least 2 days. The tissue was embedded in paraffin, sectioned at $\sim 2 \mu\text{m}$ thickness and immunochemically stained, first with either 10% normal rabbit serum or a 1:10 dilution of rabbit anti-hGH serum and then with a 1:50 dilution of FITC labelled, goat anti-rabbit immunoglobulin.

hGH produced by a tissue culture cell line expressing a metallothionein-hGH fusion gene known to give functional hGH (D.M.O. and R.D.P., unpublished data). By this criterion, hGH mRNA seems to be spliced correctly by the mouse exocrine pancreas. Analysis of the same blot using an oligomer specific for pancreatic amylase mRNA demonstrates the pancreas-specific expression of this endogenous gene (Fig. 2*b*). We used amylase mRNA as a control for contamination of other tissues with pancreas (Table 1 legend) and as an internal RNA size marker.

To ascertain whether expression was specific for the exocrine cells of the pancreas, we stained sections of the pancreas with rabbit anti-hGH and fluorescein-labelled goat anti-rabbit immunoglobulin (Fig. 3). The acinar cells of mice expressing the elastase-hGH gene fluoresced brightly, whereas no fluorescence was detectable in the islets of Langerhans (arrows), lymph nodes (not shown), connective tissue and blood vessels (Fig. 3*a*) or in the pancreas of control mice (not shown). The lumen of the pancreatic ducts also fluoresced brightly (arrow, Fig. 3*b*). Substituting normal rabbit serum for anti-hGH serum resulted in no fluorescence (Fig. 3*c*). This analysis suggests that hGH,

an endocrine hormone, can be translated properly and secreted into the gut (where it is biologically ineffective) by pancreatic exocrine cells.

Cellular differentiation follows a genetically defined path leading to the formation of adult tissues. Elucidation of the molecular mechanisms regulating those genes expressed uniquely in terminally differentiated cells will provide insight into the mechanisms guiding the development of a particular tissue. Comparison of the region of the elastase I gene which is sufficient for pancreas-specific expression with comparable regions of other pancreas-specific serine protease genes reveals a conserved 25-nucleotide sequence (Fig. 1*c* and ref. 15). This sequence is included in the chymotrypsin gene constructs that are expressed specifically in pancreatic exocrine cells in culture⁴. The presence of *cis*-acting DNA sequence elements with possible enhancer-like properties^{4,15} implies that regulatory gene products may interact with these elements to control gene expression during development of a particular tissue.

We thank Mary Yagle for technical assistance, Myrna Trumbauer for DNA microinjection, Marion Turner for preparing the manuscript, and Drs Jim Clagett, Stan McKnight and Gary Striker for their help and advice with the immunofluorescence. Rabbit anti-hGH serum was provided by the National Institute of Arthritis, Diabetes, Digestive and Kidney Diseases. D.O. was supported by the NIH Medical Scientist Training Program at the University of Washington. This research was supported by the NIH.

Received 4 September; accepted 15 November 1984.

- Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717-728 (1983).
- Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729-740 (1983).
- Queen, C. & Baltimore, D. *Cell* **33**, 741-748 (1983).
- Walker, M. D., Edlund, T., Boulet, A. M. & Rutter, W. J. *Nature* **306**, 557-561 (1983).
- Brinster, R. L. *et al.* *Nature* **306**, 332-336 (1983).
- Storb, U., O'Brien, R. L., McMullen, M. D., Gollahon, K. A. & Brinster, R. L. *Nature* **310**, 238-241 (1984).
- Swift, G. H., Hammer, R. E., MacDonald, R. J. & Brinster, R. L. *Cell* **38**, 639-646 (1984).
- McKnight, G. S., Hammer, R. E., Kuenzel, E. A. & Brinster, R. L. *Cell* **34**, 335-341 (1983).
- Palmiter, R. D., Norstedt, G., Gelinas, R. E., Hammer, R. E. & Brinster, R. L. *Science* **222**, 809-814 (1983).
- Wagner, E. F., Stewart, T. A. & Mintz, B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5016-5020 (1981).
- Stewart, T. A., Wagner, E. F. & Mintz, B. *Science* **217**, 1046-1048 (1982).
- Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D. & Costantini, F. D. *Cell* **34**, 343-358 (1983).
- Seeburg, P. H. *DNA* **1**, 239-249 (1982).
- Hammer, R. E., Palmiter, R. D. & Brinster, R. L. *Advances in Gene Technology* Vol. 1 (eds Ahmad, F., Black, S., Schultz, J., Scott, W. A. & Whelan, W.) 52-60 (ICSU Press, Miami, 1984).
- Swift, G. H. *et al.* *J. biol. Chem.* (in the press).
- Durnam, D. M. & Palmiter, R. D. *Anal. Biochem.* **131**, 385-393 (1983).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **24**, 5294-5299 (1979).
- Ashley, P. L. & MacDonald, R. J. *Analyt. Biochem.* **140**, 95-103 (1984).
- Glisin, V., Crkvenjakov, R. & Byus, C. *Biochemistry* **13**, 2633-2637 (1974).
- Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. *Biochemistry* **16**, 4743-4751 (1977).

Corrigendum

Flexure of the continental lithosphere beneath Apennine and Carpathian foredeep basins

L. Royden & G. D. Karner

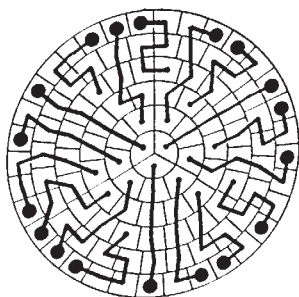
Nature **309**, 142-144 (1984).

THE vertical force (F) given in Figs 2*b* and 3*b* should be increased by a factor of two to $2.6 \times 10^{15} \text{ dyn cm}^{-1}$ and $3.0 \times 10^{15} \text{ dyn cm}^{-1}$, respectively. The text, p. 144, paragraphs 3 and 4, should be changed accordingly. Thus the applied force is equivalent to a negative buoyancy force generated by a slab with cross-sectional area of 2,600 to 3,000 km^2 in a vertical plane perpendicular to the trench axis. The curves shown in Figs 2 and 3 and the discussion remain otherwise unchanged. This does not change the fundamental results of our study, and, if anything, strengthens the arguments made in the paper—that a subsurface load plays a major role in the formation and maintenance of the Apennine and Carpathian foreland basins. Estimates of the applied vertical force ($2.6\text{--}4.0 \times 10^{15} \text{ dyn cm}^{-1}$) are now in close agreement with estimates of the vertical trench force ($\sim 5 \times 10^{15} \text{ dyn cm}^{-1}$) obtained by Davies²¹.

Molecular conformations in surfactant micelles

DILL *et al.*¹ published recently a multi-coloured representation of a 38-chain decanoate micelle constructed according to their interphase theory, and they have implied that their model differs substantially from the one we have been advancing for some time^{2,3}. As the problem of micellar structure has been fraught with controversy and even 'outlandish hypotheses'⁴, some clarification of the subject is desirable. In fact, contrary to Dill *et al.*¹, there is surprising agreement between their calculations and our model based on physical organic data^{2,3}. Among the more important features of micellar structure that we have specified repeatedly are: (1) a rough micelle surface; (2) 'fatty patches' leading to an abundance of hydrocarbon-water contact where water-insoluble materials bind hydrophobically; (3) transient water-filled cavities; (4) chain bending and looping that place terminal methyls near the surface; (5) a water-free hydrocarbon core of <50% of the micellar volume. But the picture of Dill *et al.*¹ contains all these features. Thus, theory and experiment, although perhaps not in total accord, are at least in reasonable agreement.

Our recent paper² did indeed take strong issue with the Dill-Flory lattice model of micelles⁵ presented in 1981 and quoted widely since then⁶ (see figure). The



published picture possesses: (1) a smooth surface; (2) an absence of chain looping; (3) no terminal methyl groups at the surface; (4) a completely radial distribution of chains; (5) a crystalline interior (Dill and Flory write¹: "Crowding of the chains near the core centre imposes a degree of order approaching that in a crystal.")—agreeing neither with our experimental data nor with the more recent model of Dill *et al.*¹. One cannot, however, charge Dill with inconsistency, because the Dill-Flory lattice model was not intended to represent a spherical micelle, but rather a cross-section of an infinitely-long cylinder. Unfortunately, as this picture was the only one appearing in an article devoted to micelles, many investigators mistook it for

a micelle⁷. Further use of this Dill-Flory 'micelle' will only perpetuate an irrelevant and totally unacceptable structure.

One final point of misunderstanding must be mentioned. In referring to our recent work, Dill *et al.*¹ write, 'it is claimed that micellar chains form 'loops' with termini 200 times more exposed to solvent than [are] internal segments'. But we made no such claim. Indeed, we specifically wrote², "All that can be stated with certainty is that internal olefin is less exposed to water than is the thoroughly-wet terminal olefin" and the value of 200 cited by Dill *et al.* was deduced incorrectly from our rate data. Incidentally, these rate data, pertain to the oxidation of double bonds at various locations along the surfactant chains. Comparisons of reactivity among the olefinic surfactants are legitimate because non-ideal mixing in the micelle, if it exists (it has been suggested⁸ that a "terminal double bond does not display a significant tendency to act as a second headgroup"), would remain constant in the series.

The recent model of Dill *et al.*¹ does not take into account head-group interactions, counter-ion binding, micellar water and deviations from spherical geometries. Yet even at this early stage of understanding, it is encouraging that theory and experiment are beginning to merge.

F. M. MENDER

Department of Chemistry,
Emory University,
Atlanta, Georgia 30322, USA

1. Dill, K. A. *et al.* *Nature* **309**, 42-45 (1984).
2. Menger, F. M. & Doll, D. W. *J. Am. chem. Soc.* **106**, 1109-1113 (1984).
3. Menger, F. M. *Surfactants in Solution* Vol. 1 (eds Mittal, K. L. & Lindman, B.) 347-358 (Plenum, New York, 1984).
4. Butcher, J. A. & Lamb, G. W. *J. Am. chem. Soc.* **106**, 1217-1220 (1984).
5. Dill, K. A. & Flory, P. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 676-680 (1981).
6. Fendler, J. H. *Membrane Mimetic Chemistry*, 9-15 (Wiley, New York, 1982).
7. Bendedouch, D., Chen, S. H. & Koehler, W. C. *J. Phys. Chem.* **87**, 153-159 (1983).
8. Sprague, E. D., Dueker, D. C. & Larrabee, D. C. *J. Colloid Interface Sci.* **92**, 416-421 (1983).

DILL REPLIES—Although the present state of agreement between the theory of micellar organization and experiment is indeed most encouraging, it should be noted that many of the conceptual paradoxes, which have long plagued the subject of micellar structure, have not been resolved by intuitive models^{1,2}; more quantitative treatment has been required. In comparing his views to our interphase theory^{3,4}, Menger cites five points of agreement and five of disagreement. These comparisons, however, are inappropriate in two respects: (1) only our figures and diagrams are at issue, and not the underlying theory; (2) several of his claims are inaccurate.

For example, Menger has concluded from our Fig. 3 (ref. 3) that micelles should have rough surfaces, but from our Fig. 1 (ref. 4), that they are smooth. It is not the theory which is at issue here, for the underlying model is identical in that respect. Nor are the figures contradictory; it is simply that they are intended to serve two quite different purposes. In one figure, methylene groups are represented explicitly and in the other, schematically. In the latter, irrelevant details of chain structure are neglected explicitly to demonstrate the organization of viable chain conformations and, as such, to draw inferences about surface roughness from such a diagram is inappropriate.

In spite of the potential for misinterpretation, the figure with which Menger takes issue, and which is reproduced above, serves a most useful function, showing directly how chains can fully pack space and sustain disorder. That our figure represents cylindrical micelles as opposed to spherical ones is quite unimportant and, as Menger observes, it is appropriately labelled in any case. Indeed, his claim that others have mistaken its meaning is not supported by his reference to the work of Bendedouch *et al.*⁵. Although even the more realistic computer graphics Fig. 3 (ref. 3) oversimplifies significantly certain aspects of micellar structure (shape fluctuations of individual micelles undoubtedly occur on a rapid time scale), the assumption that micelles can be represented by averaged geometric shapes, such as spheres or cylinders, is of enormous utility, particularly for the description of properties which are averages over long times and over the large number of aggregates in solution.

Some of the implications of Menger's contribution are inaccurate. It is thus worth reiterating the principal features of micellar structure observed experimentally they are consistent with the premises and predictions of the interphase model. (1) The micellar core is virtually devoid of water^{3,5}. It is wrong to imply that there are significant amounts of 'water-filled cavities' and internal 'micellar water'. Moreover, 'transient cavities' is a term which has yet to benefit from definition and we are unaware of any experiments which have a bearing on their existence. (2) Micellar chains are highly disordered^{3,4,6}. Neither theory nor experiment can be construed to imply that the micellar core is crystalline or that the chains are radially aligned. (3) Chain bending and disorder are responsible for the distribution of chain ends near the micellar surface, but to varying degrees depending on chain length and micellar geometry. In general, the shorter the chains and the smaller the micelle, the greater the incidence of ends at the surface.

As noted previously³, we disagree with the conclusions of Menger and Doll². They claim chain termini are significantly more exposed to water than are internal chain segments, stating² "[Our results are] interpreted in terms of coiling and disorder which place chain termini in the water-rich Stern region . . . The data in Table I prove that the chain termini are exposed to water a large proportion of their lifetime in micelles . . . Our data (along with those of Breslow) thus lead collectively to an important conclusion: chain bending and micelle disorder place the majority of termini outside the micelle core and within the water-rich Stern region". We believe that their caveat, however, is more accurate: the ratio of exposure of internal and external groups cannot be assumed to be equal to the ratio of oxidation rates, and thus no quantitative conclusion can be drawn from their experiments. A small degree of enhanced terminal exposure is to be expected⁷ and has been found in other experiments, including those of Breslow⁸, as terminal methyls occupy about twice the volume of internal methylenes and may be more polarizable. Factors greater than this, however, are unlikely in unperturbed micelles and would have been detected in the neutron scattering experiments we have reported³. What is at issue is whether the experiments of Menger and Doll have probed the structure of unperturbed micelles and, as we noted previously³, there is little evidence that they have, irrespective of the issue of nonideality of mixing, for the locations of the substantially-different probes used to report the accessibilities of internal and terminal groups are unknown. Their data show only that probe chains of different lengths have different oxidation rates; that is, that 15, 16 or 18 carbon chains, containing a double bond near the middle, oxidize more slowly in micellar solutions than 11-carbon chains containing a terminal double bond. Experiments dependent on the use of perturbing probes must be interpreted with caution. On the other hand, the development of nonperturbing methods such as nuclear magnetic resonance and neutron scattering has contributed enormously to our understanding of micellar chain organization.

KEN A. DILL

Departments of Pharmaceutical
Chemistry and Pharmacy,
University of California,
San Francisco, California 94143, USA

Sex and frequency of gene conversions in meiosis

LOH and Baltimore recently made the interesting observation that apparent gene conversion events in mouse major histocompatibility complex (MHC) genes occur much more frequently during female than male gametogenesis¹. Indeed, in 16 cases for which the origin of the mutated allele could be ascertained because the mutation was detected in an F₁ animal, all the mutations had arisen on chromosome 17 contributed through the egg. All 16 mutations are of the gain/loss type which, as Loh and Baltimore explain, are the most likely to have involved a gene conversion-like event. Here I suggest that this fact can be ascribed to the difference in duration of meiosis between males and females rather than an exclusive female germ-line-specific chromosome pairing/recombination mechanism.

In female mice, meiosis lasts from day 14 of gestation until ovulation because female germ cells arrest in meiotic prophase at the dictyate stage²⁻⁴; this stage can be as long as the reproductive lifespan of the animal, that is, at least 14 months. Meiosis in male mice is constant and shorter (~12.5 days)^{2,5}. Female mice are sexually mature at 8 weeks old, thus the chromosomes present in their first batch of ovulated eggs will have been paired with their homologues in meiotic prophase for 61 days. Similarly, eggs from 20-week-old females will contain chromosomes which have undergone meiosis for 145 days, and the chromosomes of eggs from 14-month-old females will have been in meiosis for 430 days. Therefore, for these three ages, meiosis is 4.9, 11.6 and 34.4 times longer than in males.

If one assumes that gene conversion occurs throughout meiosis and at the same rate in males as in females, the number of mutations observed should be directly proportional to the duration of meiosis. Hence, the probability that in a sample of 16 mutants, all 16 gene conversion events occurred in female gametes is given by the binomial $P = (r/1+r)^{16}$ where $r =$ duration of female meiosis/duration of male meiosis. The probabilities of this outcome for the three age groups of mice described above are 0.05, 0.27 and 0.63, respectively. As the ages of the females in whose gametes the mutations occurred are not available in the references cited by Loh and Baltimore, although one of the mutational screens was carried out at the Jackson Laboratories where BALB/c breeding females may be 10 months old (E. Sargent, personal communication) it is not possible to reach a conclusion as to the statistical significance of the data. It is apparent, however, that part of the reason for gene conversion being more frequent in female than male mice could be the increased duration of female meiosis. As a test for this notion, it could be determined whether gene conversion

events occur more often in mice derived from old females, that is, from old gametes.

As it is unclear what pairing conditions, if any, or chromosome associations are required for gene conversion to occur during meiosis, there is no reason to suppose that these mechanisms are drastically different in the two sexes, although an approximately two-fold difference in the intra-H-2 recombination rate is observed between males and females. That the conditions are different from those required for recombination may be inferred from Chovnick⁶, who showed that conversion events without flanking marker exchange at the *Rosy* locus in the fruitfly *Drosophila melanogaster* are of the same order of magnitude in a paracentric inversion heterozygote and a normal heterozygote, although cross-over in the inversion heterozygote was strongly suppressed, as expected. Recent experiments on mitotic and meiotic gene conversion in yeast have led to the suggestion that conversion and recombination are fundamentally different events⁷. Chovnick's experiment is relevant to *Mus musculus* because MHC gene evolution has occurred over the past 1 Myr in different mouse complete *t*-haplotypes^{8,9}, which are naturally occurring inversion heterozygotes for the MHC¹⁰. Crossing-over with wild-type chromosomes is suppressed but it is possible that gene conversion by intrachromosomal genetic transfer from wild-type chromosomes or interchromosomal transfer has been a mechanism partly responsible for their evolution.

Finally, if female-specific gene conversion is an effective means of generating genetic polymorphism in addition to that obtained by recombination, and if it also occurs at multigene loci other than the MHC and in other organisms, then it may contribute to the successful evolution of two mechanisms for generating gametes.

KEITH R. WILLISON

Institute of Cancer Research,
Chester Beatty Laboratories,
Fulham Road, London SW3 6JB, UK

1. Loh, D. Y. & Baltimore, D. *Nature* **309**, 639-640 (1984).
2. Lyon, M. F. *Mutat. Res.* **87**, 323-345 (1981).
3. Speed, R. M. *Chromosoma* **85**, 427-437 (1982).
4. Dietrich, A. J. J. & Mulder, R. J. P. *Chromosoma* **88**, 377-385 (1983).
5. Oakberg, E. F. *Nature* **180**, 1137-1138 (1957).
6. Chovnick, A. *Genetics* **75**, 123-131 (1973).
7. Fink, G. R. & Petes, T. D. *Nature* **310**, 728-729 (1984).
8. Nizetic, D., Figueroa, F. & Klein, J. *Immunogenetics* **19**, 311-320 (1984).
9. Dembic, Z., Singer, P. A. & Klein, J. *EMBO J.* **3**, 1647-1654 (1984).
10. Shin, H.-S. *et al.* *Nature* **306**, 380-383 (1983).

1. Menger, F. M. *Acc. chem. Res.* **12**, 111-117 (1979).
2. Menger, F. M. & Doll, D. W. *J. Am. chem. Soc.* **106**, 1109-1113 (1984).
3. Dill, K. A. *et al.* *Nature* **309**, 42-45 (1984).
4. Dill, K. A. & Flory, P. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 676-680 (1981).
5. Bendedouch, D., Chen, S. H. & Koehler, W. C. *J. phys. Chem.* **87**, 153-159 (1983).
6. Dill, K. A. *J. phys. Chem.* **86**, 1498-1500 (1982).
7. Gruen, D. W. R. *J. Colloid. Interface Sci.* **84**, 281-283 (1981).
8. Breslow, R., Kitabatake, S. & Rothbard, J. *J. Am. chem. Soc.* **100**, 8156 (1978).

Gardner's choice

Walter Gratzer

The Sacred Beetle and Other Great Essays in Science.

Edited by Martin Gardner.

Prometheus, 700 East Amherst St, Buffalo, NY: 1984. Pp.427. \$22.95.

To be published in Britain later this year by Oxford University Press.

A PART of the enjoyment that an anthology affords is to stimulate the reader's prejudices, especially as to what has been omitted. My first reaction to Martin Gardner's collection of essays was regret that modesty should have deterred him from including some of the most entertaining of modern science writing, to be found in his own books, such as *In the Name of Science* and *Science, Good, Bad and Bogus*. Other notable omissions are J.B.S. Haldane, whose entire *oeuvre* I should have been tempted to retain, and Medawar, as well as the English natural historians, such as Richard Jefferies and Gilbert White. There are, however, many pleasurable surprises in this highly individual selection — a somewhat enlarged hardback version of a paperback published in 1957 (which is no doubt why the late Samuel Goudsmit, whose dates are correctly given as 1902–1978, is stated to be the current chairman of the Physics Department at Brookhaven).

Gardner has disinterred some late nineteenth century romantics, such as John Burroughs, throbbing with Victorian fervour and sincerity, and (to me at least) a rather glutinous read. There are some choice examples of the thinking of Freud and of Havelock Ellis, which could be said to stretch the definition of science. Attend: ... fair hair is more beautiful because it harmonises better with the soft outlines of woman, ... the hair of the armpits, also, Stratz considers, should be light. On the other hand, the pubic hair should be dark in order to emphasize the breadth of the pelvis and the obliquity of the angle between the mons veneris and the thighs.

And so on for pages. (If one suspected Havelock Ellis of the possession of a sense of humour one might well wonder whether he had not invented the unlikely Stratz — perhaps the original inspiration for Dr Strabismus of Utrecht, whom God preserve.) Then we are assailed by Maeterlinck on the bee, apostrophizing the reader in remorselessly purple prose. (How incomparably better Miriam Rothschild.) Fabre, on the other hand, is irresistible, his slightly dotty enchantment with the habits of the dung beetle a constant delight.

Julian Huxley's dissertation on the courtship of birds seems oddly dated, with its patina of anthropomorphism. The crested grebe, he avers, though constrained by the shackles of monogamy, flirts, when frustrated, with passing ladies; his consort, who until then has remained indifferent to his advances, thereupon attacks and drives

off her rival, and only after this is there a loving reconciliation. "How human", Huxley muses, "again we see to what a pitch of complexity the bird's emotional life is tuned". Well perhaps, though to me it only recalls the cartoon that shows the zoo gorilla bearing away a shrieking blonde; her husband, observing with composure, suggests: "Why don't you tell him you've got a headache?"

There are some well-chosen contributions by Darwin and T.H. Huxley, whose majestic cadences retain their power to hold the reader in thrall. Stephen Jay Gould makes his appearance in the new edition (in the company of Asimov, Sagan and Lewis Thomas), with a beautifully wrought essay on the amorality of nature and the contortions of the Victorian moralists in the face of such *Grand Guignol* as the behaviour of the ichneumon fly, which injects its eggs into a caterpillar, so that the larvae, when they hatch, devour the still living but paralysed host from within. The best that a leading entomologist and divine of the day could manage, to justify God's ways to man, was to praise the maternal regard of the wasp in undertaking its hazardous operation. (His predicament has been encapsulated in the words of one of our latter-day gurus: the leopard will lie down with the kid, but the kid won't get much sleep.)

Among the more surprising appearances in *The Sacred Beetle* are those of G.K. Chesterton, Robert Louis Stevenson, John dos Passos and a science editor of *Time* magazine, Jonathan Norton Leonard.

There is also A.N. Whitehead, seeking to reconcile science and religion, and vexing Galileo's ghost with an ingenious rehabilitation of pre-Galilean astronomy by recourse to relativity. Ortega y Gasset inveighs against science, which he evidently feared would destroy the values of European civilization by spawning a new breed of insufferable and dangerous "learned ignoramuses", too narrow in education and vision to recognize the limitations of their competence. This is a sustained piece of polemical virtuosity, and one cannot help but wonder how Ortega, a pillar of the Spanish Republic, got along with his Prime Minister and fellow professor, Juan Negrín, a physiologist, who chose to organize a scientific meeting in Madrid at the height of the Civil War with the enemy already at the very gates. Did he even perhaps have Negrín in mind?

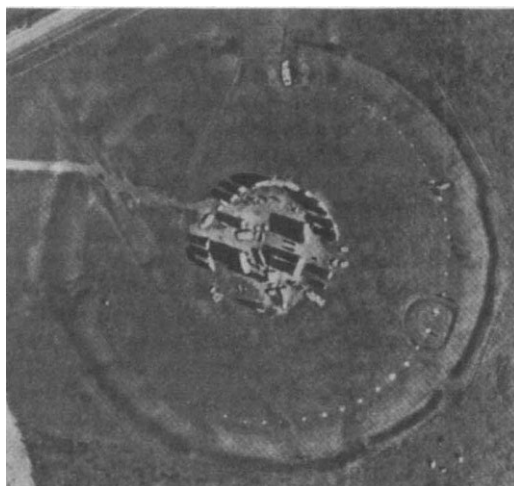
Among other pleasures that Gardner puts in our way in this odd but engaging book are Goudsmit's account of what he found when he entered Germany with the US Army in 1945 to investigate how science had fared under Hitler; a dazzling philippic on determinism by Eddington; two essays by Bertrand Russell, pellucid in substance as in style; and H.G. Wells, as compelling now as then in a chapter from one of the scientific romances. Here he is in a passage that Gardner has quarried from an almost forgotten novel, *Meanwhile*:

The disease of cancer will be banished from life by calm, unhurrying, persistent men and woman ... and the motive that will conquer cancer will not be pity nor horror; it will be curiosity to know how and why. 'And the desire for service', said Lord Tamar. 'As the justification for that curiosity', said Mr Sempack 'but not as the motive. Pity never made a good doctor, love never made a good poet. Desire for service never made a discovery'.

Well said, H.G.!

□

Walter Gratzer is in the Medical Research Council's Cell Biophysics Unit, King's College, University of London.



Stonehenge from the air, showing the outer ditch and bank as well as the standing stones themselves. The photograph is reproduced from *Prehistoric Europe* by Timothy Champion, Clive Gamble, Stephen Shennan and Alasdair Whittle. Publisher is Academic Press. The book will be reviewed in *Nature's* textbook supplement, which will appear in the issue of 7 March.

A mind ever on the stretch

Robert Siegfried

Thomas Beddoes MD 1760–1808:

Chemist, Physician, Democrat.

By Dorothy A. Stansfield.

Reidel: 1984. Pp. 306. Dfl. 140, £35.75.

THOMAS Beddoes was not a great scientist. Nor did he occupy a public position that makes a biography obligatory. History has recorded his existence chiefly in the lives of Humphry Davy and Samuel Taylor Coleridge, both of whom benefited when young from his support and stimulation. But the range of Beddoes' active involvements was so great that his biography provides a fascinating perspective on major social, political and medical issues of late eighteenth century Britain. One obituary writer characterized him as the "physician whose mind was ever on the stretch". In the present work Mrs Stansfield has described all of Beddoes' activities with the aim "to see him whole".

A native of Shropshire, Beddoes was educated chiefly at Oxford, though he took his MD there in 1786 only after two years at Edinburgh. He stayed on as Reader in Chemistry until 1792. During those years he visited Lavoisier and de Morveau in France and became an early advocate of the "new chemistry". He wrote on the comparison of basalt and granite, finding evidence to support the Huttonian plutonic geology. Having taught himself Italian, French and German, he translated Spallanzani's *Natural History* (two editions), two major works of Torbern Bergman and Carl Scheele's *Chemical Essays*.

Beddoes moved from Oxford after his radical political views had made a chemical professorship unlikely. He quickly established a successful medical practice in Bristol and began planning the Pneumatic Medical Institution for the investigation of an aerial cure for pulmonary consumption. He was able to enlist the support of many prominent men, notably Thomas Wedgwood and James Watt, both of whom had personal reasons to hope for his success. The Institution opened its doors in 1798 with the young Humphry Davy as superintendent of the laboratory. The physiological properties of nitrous oxide (laughing gas) were the only discoveries of any note, but Beddoes' failure to exploit the gas's anaesthetic effect remains puzzling.

Beddoes was also busy writing tracts supporting a wide variety of liberal and progressive causes. He collected signatures petitioning for the abolition of the slave trade; he denounced with sometimes incoherent fervour government policies regarding France, and declared the war an unnecessary evil. More temperately he

advocated equal educational opportunities for girls and boys, and wrote on the design of educational toys. After Davy's departure for London in 1801 he dedicated the Pneumatic Institution to the cause of preventative medicine, writing on the evils of drink, and producing medical guides and dietary advice for the poor.

In concentrating on Beddoes' own work, Mrs Stansfield has sometimes left the reader with an inadequate historical context by which to judge the originality or the uniqueness of Beddoes' contributions. But there is no uncertainty about his profound sense of humanity, and I am sorry that Mrs Stansfield did not make that quality the integrative theme of her biography. It could explain his failures as well as his ambitions, for he loved

humanity not wisely but too well.

There has been no major account of Beddoes since J.E. Stock's protective *Memoirs* of 1811. The usefulness of Mrs Stansfield's balanced biography is enhanced by extensive end-notes and a bibliography (somewhat idiosyncratic) of primary and secondary sources. There are two appendices; the list of Beddoes sixty-five contributions to the *Monthly Review* between 1793 and 1801 is a most impressive testament to the range of his competency. Name and subject indexes complete this study of a man who saw so much more need in the world than he could meet. □

Robert Siegfried is Professor of the History of Science at the University of Wisconsin-Madison.

Cosmic magnetism

Leon Mestel

Magnetic Fields in Astrophysics.

By Ya. B. Zeldovich, A.A. Ruzmaikin and D.D. Sokoloff.

Gordon & Breach: 1984. Pp.365. \$80.

MAGNETOHYDRODYNAMICS — the study of the two-way interaction between the kinetic and magnetic fields in electrically conducting fluids — grew out of cosmical observations, beginning with the variations in the geomagnetic field, and the variety of phenomena grouped together as "solar activity". It is the large cosmical length-scales which facilitate a strong interaction, for by Ampère's law a large-scale field is maintained by currents that are comparatively weak and so suffer far less Ohmic dissipation than in laboratory experiments. The subject is now an autonomous branch of fluid mechanics, but it is in the astrophysical context that one still finds the greatest variety of application.

Probably the most fundamental question is that of the origin and maintenance of cosmical magnetic fields. After a long period of doubt as to whether "dynamo action" was possible (due largely to results such as Cowling's "anti-dynamo theorem", which ruled out the simplest and most easily pictured field structures), there is now almost a plethora of models with a large-scale magnetic field that grows spontaneously via motional induction.

The present volume contains an extensive discussion of the conditions for dynamo action, starting with Faraday's unipolar inductor, crucially modified so that the currents induced in the rotating disk complete their circuit through a fixed helical wire, twisted in the same sense as the rotation. This simple model contains the two elements — differential rotation and helicity — that are typical of hydromagnetic dynamos (as anticipated in a seminal paper by E.N. Parker). The authors show considerable pedagogic skill

in bringing out both the topological and analytical aspects of the problems; for example, they contrast the limitations of two-dimensional flows, which allow only "slow" dynamos, developing in time only by virtue of finite resistivity, with the greater potentiality of three-dimensional, which in principle can yield "fast" dynamos that continue to operate at infinite magnetic Reynolds number.

Limitations on length have made inevitable a less detailed mathematical discussion of specific astrophysical problems, even of those that are direct applications of dynamo theory such as the cyclical magnetic activity shown by the Sun and detected in solar-type stars ("the solar-stellar connection"). Still, the wide-ranging astrophysical reader will certainly find very useful the inclusion in one volume of chapters on the observational evidence for cosmical magnetism, and in particular for the galactic field, on magnetism and star formation, magnetic fields in cosmology, magneto-accretion by black holes, and on strong magnetic fields emanating from neutron stars.

The book remains to some extent personal in its emphasis — for example the authors say nothing about hydromagnetic stability, and little about the problems of the strongly magnetic early-type stars. There are also signs of somewhat hasty preparation in the occasional repetitiveness and premature reference to difficult concepts discussed subsequently. Nevertheless this is a very valuable addition to the literature, and one hopes that the publishers will soon offer the book in a more reasonably priced edition. □

Leon Mestel is Professor of Astronomy at the University of Sussex.

● Peter and Jean Medawar's *Aristotle to Zoos: A Philosophical Dictionary of Biology* has now appeared in paperback. British publisher is Oxford University Press, price is £4.95. In the United States the paperback edition will be published in March by Harvard University Press. For review see *Nature* 308, 800 (1984).

Penrose's way with relativity

Abhay Ashtekar

Spinors and Space-Time. Vol. 1 Two-Spinor Calculus and Relativistic Fields.

By R. Penrose and W. Rindler.

Cambridge University Press: 1984.

Pp. 458. £45, \$89.50.

ALTHOUGH spinors were introduced by Elie Cartan about 70 years ago, their use in general relativity became common only rather recently. It is in fact easy to trace the history of this development.

The first major event was the lecture by Roger Penrose, entitled "General Relativity in Spinor Form", given at the third international conference on general relativity and gravitation, held at Royaumont, near Paris, in June 1959. The reason behind the success of spinor methods is perhaps best explained in the opening sentence of Penrose's report in the proceedings of the conference. It reads:

I propose to show that if a spinor calculus is used in general relativity, instead of the usual tensor calculus, many of the important expressions of the theory take on a surprisingly simple form.

The elegance and the power of the new techniques made an immediate impact. I am told that a group of younger participants, several of whom were destined to become leaders in general relativity, sought out Penrose for additional evening lectures during the conference itself. Spinorial techniques soon began to appear in numerous research articles, particularly in the subject of gravitational radiation.

The field next received a boost from the detailed article by Felix Pirani in the proceedings of the 1964 Brandeis summer school, which made the subject accessible to beginning research students.

Penrose himself had given a series of lectures on spinors somewhat earlier and the present monograph grew out of notes taken by Wolfgang Rindler. Although the existence of these notes was well known in the relativity community, they did not receive wide circulation and their appearance in book form, polished and updated, has been eagerly awaited for more than 15 years! During the hiatus the field has grown and with it the size of the book. It is now a work in two parts. The second volume, to be published at the end of this year, will contain material on null infinity and twistor theory, as well as certain recent developments such as the proof of the positive energy theorems and the definition of quasi-local conserved quantities.

Roughly, one can think of spinors as square roots of vectors; it takes the product of a spinor with its complex conjugate to make a vector. It is therefore only natural to regard spinors as the primitive objects from which vectors and tensors can be built. In this, though, there is a price to be

paid. While one needs only a differential manifold to speak of tensor fields, one needs a manifold with a metric to introduce spinors. However, in general relativity, space-time manifolds do, naturally, come with metrics which dictate physical processes such as propagation of light and ticking of clocks. In this setting, conceptually, spinors are indeed a more natural starting point than world-vectors and tensors. The use of spinors also has technical advantages. Since the spin-space is a two-dimensional (complex) vector space, it comes equipped with a unique (up to an overall multiplicative factor) non-degenerate 2-form. Together with its inverse, the 2-form can be used to identify contravariant and covariant spinors, that is, to raise and lower indices. Further, a spinor which is anti-symmetric in any two of its indices is proportional, in those indices, to the 2-form and the proportionality factor is easily obtained by contracting the two indices in question; consequently only symmetric spinors matter. These simplifications are not shared by world-tensors because the tangent space at a space-time point is four (real) dimensional.

The monograph first introduces the reader to the geometry of spinors in Minkowski space, then treats in detail spinor algebra and calculus in the abstract index notation and concludes by applying the framework to the study of various physical fields in curved space-time. It is written with a wonderful sense of balance. Thus, while the treatment is clearly rigorous, the style is not terse. There are remarks to assist the reader to see why one is doing what one is doing and where one is headed, and beautiful figures to help visualization of many constructions. Consequently, the book will serve at least three purposes: it can be used as a text in a "special topics" course in general relativity or mathematical physics; it can be read profitably by individual graduate students and beginning researchers wanting to become acquainted with the spinorial description of the world on their own; and it will be a very valuable source for advanced researchers. Here, at last, one can find, in one place, the numerous spinorial formulae — not to mention constructions, theorems and their proofs — that researchers in general relativity so often need, presented in the most general context.

In addition, the book has several novel features which will make it attractive to a wider variety of readers. One is the systematic development and use of the abstract index notation. This notation was invented by Penrose in order to treat tensor fields in an intrinsic, coordinate-free way and yet retain the flexibility offered by the presence of indices, which is so useful in long calculations. In this notation, indices serve simply as abstract markers which tell the nature of the tensor considered (for example its contravariant versus covariant character) and do not take on numerical

Advertisement



Keep in touch with the foremost laser research through

LASER CHEMISTRY

AN INTERNATIONAL JOURNAL

Editors: *Asia* Koichi Shimoda, Keio University, Japan. *East Europe & USSR* V. S. Letokhov, USSR Academy of Sciences. *Western Europe* J. Jousset-Dubien, Laboratoire de Chimie Physique, Talence, France & R. J. Vetter, CNRS II, Orsay, France. *USA & Canada* Paul J. Dagdigian, Johns Hopkins University & Ahmed H. Zewail, California Institute of Technology, USA

"... with the ferment of activity in laser research and the quality of its international board, this new journal should rapidly gain momentum and respectability" — The Times Higher Education Supplement

LASER CHEMISTRY is the first journal to unite international research activity in this rapidly developing field. It is intended to bridge the gap between physics and chemistry laser-related research. The journal publishes original theoretical and applied papers as well as overviews covering a large number of topics, but with particular focus on the following: laser-selective chemistry; reactive scattering and state-to-state dynamics; multiphoton dissociation, ionization and absorption phenomena; laser-assisted collisions and reactions in buks and on surfaces; primary processes in photochemistry and chemiluminescence; molecular relaxation processes; dephasing, energy redistribution and ultrafast dynamics; new laser techniques; application of lasers in biology.

For speedy publication, **LASER CHEMISTRY** also publishes 'rapid communications' in camera-ready form.

Subscription Information:

6 issues per volume ISSN: 0278-6273
Current volume: Volume 6, 1984-1985
Institutional rate: \$146/volume
Individual rate: \$74/volume

(Individual rates available only to those who subscribe directly from the publisher and pay through personal cheques or credit cards.)

SEND FOR FREE SAMPLE COPY

To order, send cheque or credit card information to:



**Harwood
Academic
Publishers**

Marketing Dept, 50 West 23 Street,
New York, NY 10010, USA or 1 Bedford
Street, London WC2E 9HD, UK

values. Mathematicians should find this notation most interesting. Another feature is that 2-component spinors are used throughout the book; the simplicity and utility of these objects is thus brought home and high energy physicists may well succumb to their charm. Also, the definition of a manifold in terms of its ring of smooth functions should please many a mathematical physicist. While this approach is not new, to my knowledge this is the first time that it appears in a monograph accessible to physicists.

There are, however, a few omissions that I personally regret. The first is the geometrical interpretation of Lie derivatives, and of curvature tensors. While the "algebraic or axiomatic approach" to derivative operators used here is excellent, in that it brings to the forefront the properties that one frequently needs in practice, graduate students would have been greatly helped if the geometrical structure involved had been somewhere explained. Secondly, it would have been helpful if a symbol — the abstract index analogue of the Infeld–Van der Waerden symbol — representing the isomorphism between hermitean spinors and tensors had been first introduced explicitly and then dropped in calculations involving a fixed space-time metric. When one deals with a fixed conformal class of metrics, this symbol can be dispensed with and suitable spinor indices can be simply identified with tensor indices. However, sometimes — for instance in certain perturbation calculations as well as in quantum gravity — one needs to work with a wider class of metrics. In these cases, it is important to keep track of how the chosen isomorphism between spinors and tensors changes with the choice of the metric, particularly because the metric does not determine the isomorphism uniquely. An inexperienced reader may miss this point altogether. Finally, I feel that it would have been useful to have a section or an appendix summarizing the relation between the 2-component spinors and the 4-component ones.

These points by no means detract from the stature of the work as a whole. In recent developments in classical gravity, there have been instances of constructions which are essentially spinorial in the sense that they cannot even be recast in the language of world-tensors. There are growing indications that such constructions may also play an important role in quantum gravity. Thus, the book has come out just in the wake of the realization that spinors may play an even more profound role in theoretical physics than they do today. Because of this, and because Penrose's ideas have influenced the younger generation of relativists so deeply, it is very likely to become a classic. □

Abhay Ashtekar is a Professor of Physics at Syracuse University, New York, and at the University of Paris VI.

Pronouncing on the language

Joanna Czechowska

The English Language.

By Robert Burchfield.

Oxford University Press: 1985. Pp.194.
£9.50, \$19.95.

DURING the fifth century AD the Roman establishment moved out of the British Isles and Germanic tribes from Northern Germany and southern Denmark moved in. The descendant of the language they spoke is now the mother tongue of more than 300 million people worldwide and the second language of many millions more. Robert Burchfield's book documents the history of English from the fifth century to the present day, tracing the changes caused by the absorption of other languages, the influence of English settlers in overseas colonies and the increasing introduction of new words designed to cope with technological advances.

Old English, spoken from the fifth century until the Norman Conquest, was an inflected language belonging to the Germanic family. It was written down using runic characters, an alphabet consisting of linear strokes which could easily be carved onto wood or stone. Even when the Roman alphabet was adopted, certain runic characters remained for a time to symbolize peculiarly English sounds. The advent of Christianity in the sixth century brought about the introduction of many Latin words, but it was after the Conquest of 1066 that the language changed dramatically when Norman French words became absorbed, leaving traces which are still discernible.

The introduction of printing in the fifteenth century fossilized the spelling of English while pronunciation continued to change. This accounts for the considerable discrepancy in the way some English words are spelt and the way they are pronounced. British settlers in the United States, Canada, Australia, New Zealand and South Africa adapted the language to describe new features they encountered in the environment, often by borrowing words used by the native inhabitants, for example *skunk*, *kangaroo*, *moccasin*.

Dr Burchfield's summary of English is perfectly suited for anyone with some interest in the subject. It is not a specialist book but it leaves the reader with a sound overall view of the language without having to wade through the minutiae of syntax (indeed he castigates "despotic professors of linguistics [who] have failed to notice that they have taken the subject beyond the reach of intelligent laymen"). There is a fascinating section on slang — rhyming slang (*butcher's hook*, *look*; *Todd Sloan*, *alone*), back-slang (*yob*) and Citizens' Band Radio slang which has infiltrated

British English from North America. The popular myth that the English spoken in the United States is antiquated and therefore somehow "purer" than that spoken in Britain, he largely dispels; in the seventeenth century words such as *fertile*, *leisure* and *dormitory* were pronounced in both the British English and American English manner.

This is a brief but clear and immensely readable book. Fifteen-hundred years of English has been condensed into some 200 pages of text without losing sight of how rich, diverse and complex our language is. □

Joanna Czechowska is on the staff of Nature.

African past

A.C. Hamilton

Southern African Prehistory and Paleoenvironments.

Edited by Richard G. Klein.

Balkema: 1984. Pp.404. Dfl.70,
£17.50, \$24.

EVER since the discovery of australopithecines in the 1920s and 1930s southern Africa has been a focus of attention for the study of the evolution of man. The archaeological record is one of the longest in the world. Progress has been hampered by the difficulty of dating the major sites, although, as Rightmire points out in this book, some control is possible through comparison of fossil hominids and vertebrate assemblages with the better dated East African material. The Middle Pleistocene fossil hominid record of the region remains very poor.

A difficulty confronting prehistorians and palaeoenvironmentalists working in southern Africa is that few people outside their own disciplines are interested in the details of their working techniques, despite the widespread popular interest in early man. This book is aimed at the specialist and all of the articles are marked by an excellent critical approach to evaluation of the evidence. Most contributors do, however, attempt to provide a story-line and some chapters should be understandable to a wider audience. Let us hope that this is so. The various branches of Quaternary science have much to contribute to one another and, outside scientific circles, the study of human evolution is an especially important subject in the Republic of South Africa. Apartheid is buttressed partly by theological dogma which is at complete variance with estab-

● Recently published by Cambridge University Press in their *World Archaeology* series is *African Archaeology*, by David W. Phillipson, a concise, non-technical account of the subject. The book will be reviewed in a future issue of *Nature*.

lished scientific views. Bearing this in mind, the authors should be encouraged to produce another but more popular book for wider dissemination.

The number of contributors (eight) is much reduced from the companion volume published by Balkema, *The Sahara and the Nile*. This, combined with much cross-referencing and first-class editing, has resulted in a coherent and readable volume.

In his chapter, Butzer is critical of widely-embracing theories of environmental (especially climatic) change which have been advanced on the basis of isolated analyses of sediment sections, such as "a deep-sea core or a pollen profile" (p.1). This is not altogether fair and indeed, despite the geographical distance of the evidence, Volman employs oxygen-isotope stages defined in deep-sea cores as a chronological framework for discussing early hominid development. In any event the motivation for detailed studies and wider interest in the results comes from more general speculations and there can be benefit in less cautious scientists publishing what might be regarded as their after-dinner thoughts. Careful, methodical studies such as those undertaken by Butzer can be used to test or modify such hypotheses.

Scott (on pollen), Klein (Late Pliocene to Recent mammals) and others all draw extensively on modern analyses to help interpret the past. The climate of southern Africa has not generally been favourable for the preservation of pollen, but nevertheless some sites have provided useful results. Indeed, two generalizations which emerge from the book are, on the one hand, the value of good sites and, on the other, their rarity in southern Africa. An example of such a site in the field of vertebrate palaeontology is the very rich Early Pliocene Varswater Formation described by Hendey. This dates to just before the arrival of the hominids whose activities are believed to have contributed towards faunal impoverishment.

The hunter-gatherer-fisher way of life persisted in parts of southern Africa into historical times, an unusual feature which provides justification for the disproportionate amount of the book devoted to a chapter on the Later Stone Age by Deacon. The chronology of this period, and even more so that of the Iron Age (covered by Maggs), is relatively well established, thanks to the widespread application of radiocarbon dating.

This volume is a major achievement by authors, editor and publisher. It should provide an impetus for further research in southern Africa and a source of inspiration to those interested in the prehistory and environmental history of other parts of the world. □

A.C. Hamilton is a Senior Lecturer in Environmental Science at the University of Ulster.

Water from a stone

Peter J. Smith

Inscrutable Earth: Explorations into the Science of Earth.

By Ronald B. Parker.

Charles Scribner's Sons: 1984. Pp.200. \$14.95.

IN MY ideal world, which I freely admit would possess some of the more endearing characteristics of Victorian Britain, there would be a profusion of Earth-science literature, ranging from research monographs at one end of the scale to popular volumes at the other. In the real world of the 1980s the advanced tomes and student texts are common enough, but the once-green landscape of popular geological culture has long since become a desert. Or perhaps it would be fairer to say a desert sprinkled with oases, most of which are polluted. Exclude the hack-written children's volumes and the bland series from the multinational pap publishers, and what of value remains, in the English speaking world at least, can be counted on the legs of a quartered centipede.

Despite the claim in the publisher's blurb, Oasis Parker may not have been moulded with quite the finesse of Oases McPhee and Gould; but thirsty caravan-

ners, if such there be any more, will still find there welcome relief. Fourteen shortish essays cover topics both well-worn (crustal chemical composition, dating, atmospheric oxygen, stratigraphy and so on) and under-exposed (the sodium budget, heat flow, gold mining, uniformitarianism, for example). They are stylistically simple without being dull, comprehensible to the non-scientist without being patronizing, and peppered with humour and instructive metaphor. The extended stratigraphy analogy, for example, involving papers on a desk complete with dust layers (volcanic ash) and dead flies (fossils), is simple but inspired.

There is a danger, of course, that gasping travellers will allow relief at spotting sustenance to overcome critical judgement. The waters of Oasis Parker are not entirely unadulterated; there are occasional errors, misleading oversimplifications and infelicities of language. But such minor blemishes are more than outweighed by the merits, not the least of which is a nice line in oblique viewing. The idea that, since the Earth's crust is 94% oxygen by volume, one could tout lumps of old rock as "solid oxygen with minor impurities" without falling foul of the fraud squad is an image to cherish. □

Peter J. Smith is Reader in Earth Sciences at the Open University.

On the breeze

Miriam Rothschild

Milkweed Butterflies: Their Cladistics and Biology.

By P.R. Ackery and R.I. Vane-Wright.

British Museum (Natural History)/Cornell University Press: 1984. Pp. 425. £50, \$75.

MAJOR contributions to systematics are essentially useful, but they are rarely exciting, and even less frequently are pleasurable to read. For me, *Milkweed Butterflies* has all these merits and the book will no doubt prove equally fascinating for a wide range of Lepidopterists. Besides producing a new and stimulating cladistic classification of the subfamily Danainae, the authors have assembled virtually all the existing information about the group — probably the most intriguing assemblage of butterflies in the world —, supporting it with some 1,500 references.

Vane-Wright and Ackery add greatly to the interest of the morphological descriptions by providing information concerning the behaviour of the adult in the field. Thus the flight of *Danaus ismare* is fast, high and sailing, but that of *D. affinis* slow and weak. *Danaus genutia* flies slowly close to the ground, *D. melanippus* has a skipping flight and flies to the inaccessible tops of trees if disturbed, while

the Monarch (quoting Urquhart, 1960) has five distinct flight patterns. The sequestration of alkaloids by adult butterflies is also well described and illustrated. The authors give invaluable and detailed lists of food plants, and in addition take the larval and pupal characters into consideration in their classification.

It is the combination of biology, systematics and unusually interesting zoogeography that gives this monograph its special quality. The hiving off of 11 species in the genus *Danaus* which, among other characters, share the capacity to sequester and store cardiac glycosides, is a good example of this new approach. One only hopes that others will follow the lead and that the various families in the combined Rothschild and British Museum collections will be treated in a similar manner.

In a work of such dimensions there are understandably some omissions, but it is a pity the authors missed the paper in which the sequestration and storage of cardiac glycosides was first recorded in *D. plexippus* and *D. chrysippus*, for this article also mentions the extraordinary sensation of *déjà vu* which is sometimes experienced when handling these butterflies. However, the one serious sin of omission is the lack of a subject index. □

Miriam Rothschild is an entomologist with a special interest in fleas and plant-insect interaction.

Accommodation in the Upper House

W.H. Brock

All Scientists Now: The Royal Society in the Nineteenth Century.

By Marie Boas Hall.
Cambridge University Press: 1984.
Pp.261. £25, \$49.

FOLLOWING close on the heels of several books dealing with the history of the British Association for the Advancement of Science — that peripatetic "Parliament", or more precisely, "Scientific Commons" — Dr Hall has produced a valuable account of the Royal Society — that Upper House of Science — during the nineteenth century. Such contemporary parliamentary metaphors for Britain's two most important Victorian scientific organizations are not inappropriate. Until its procedures were reformed in the 1840s, the membership of the Royal Society literally consisted of unscientific peers as well as practising scientists; the effect of reform was to ensure that the Society became a truly *scientific* House of Lords — an élite body which could be approached by governments for disinterested and expert advice.

At the beginning of the nineteenth century, the Society's President, the botanist Sir Joseph Banks, had already presided for 21 years and was to reign for another 20 until his retirement and almost simultaneous death at the age of 76 in 1820. Not surprisingly the "Banksian Learned Empire", as it has been termed, fostered natural history rather than the physical sciences, sowing discontent amongst younger unpatronized talents. (Dr Hall detects a continuing tension between what eventually became the A (physical) and B (biological) constituents of the Society during most of the century.) Banks was also autocratic: potential Fellows had to meet with his personal approval and were frequently elected for their financial, social and political standing and clout rather than for scientific fame or promise. Fellows were *ipso facto* gentlemen. Nevertheless, although the Society was packed with men who agreed with Banks, as the names of Wollaston and Davy attest, it did not lack real talent.

However, the Society could not stand apart from the liberalizing tendencies of an age in which "rotten boroughs" disappeared and hitherto disenfranchised groups of middle-class men sought an end to aristocratic privilege. Outside the Society new scientific and more democratic associations were forming in metropolis and province: the Geological Society in 1808, the Astronomical Society in 1820 (which gained Royal status in 1831) and the BAAS. While recognizing the potential threat to the Royal Society's hegemony, Dr

Hall prefers not to interpret these events in terms of the cultural politics of scientific organizations fashionable among younger historians of science, but as a slow process of inevitable accommodation in which the Royal Society exemplified the "uniquely English political process ... of changing while remaining the same".

What was to remain unchanged (and strengthened in the process of accommodation) was the Royal Society's premier position; but to maintain this, first under Humphry Davy's presidency (1820–1829) and then under that of the regal Duke of Sussex (1830–1838), the Society's oligarchical and undemocratic administrative structure had to be reformed, printed *Proceedings* and Council Minutes introduced (both in 1832), papers to be more stringently refereed by experts in the area, discussion and debate (banned since the mid-eighteenth century) reintroduced, and the Society's finances strengthened by doubling the annual subscription for Fellows to four pounds. Then, in 1847, under the guidance of the lawyer-physicist Robert Grove, the Society introduced annual elections to the Fellowship (instead of the disruptive, almost weekly, elections) and restricted new Fellows to 15 in number. Although a privileged category for princes, peers and privy counsellors was retained, and can still prove controversial, this restriction on numbers ensured that the Society retained its position as Britain's national academy of science. (Albeit in the restricted English sense of "science", for when international scholarly relations between national academies became important at the end of the century it was

necessary to create a British Academy (founded 1903) for the historians and other academics excluded from the Society.)

From mid-century, following these reforms, the role of President, while reduced administratively to the chairmanship of the Council which took the decisions, acquired the deferential trappings of constitutional monarchy and as such it became filled by illustrious figures such as Rosse, Airy, Hooker, Huxley and Kelvin. As befitted its image as the Upper House of Science, the Society did much to encourage scientific research through an honours system of medals, research grants (the Parliamentary Grant was first administered in 1849) and the diffusion of information through publications, including the *Catalogue of Scientific Papers* (1867–1925) to which all historians of science are indebted. Its dramatic support for scientific investigations on voyages of exploration to the four corners of the Earth was particularly important.

All these matters receive attention from Dr Hall, as do the Society's surprisingly convivial relationships with government throughout the period. There remains much for more detailed examination by historians, but *All Scientists Now* provides an excellent overview. Despite an over-generous tendency to see "good in all men" (even the exasperating Babbage), Dr Hall's book will be enjoyed by Fellows and Commoners alike. □

W.H. Brock is Reader in the History of Science and Director of the Victorian Studies Centre at the University of Leicester.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS